

1992

## Geographic Data in City Planning Departments: The Volume and Use Related to Advancements in Geographic Information Systems (GIS) Technology

Fawzi Saeed Kubbara  
*Portland State University*

Let us know how access to this document benefits you.

Follow this and additional works at: [http://pdxscholar.library.pdx.edu/open\\_access\\_etds](http://pdxscholar.library.pdx.edu/open_access_etds)



Part of the [Urban, Community and Regional Planning Commons](#)

---

### Recommended Citation

Kubbara, Fawzi Saeed, "Geographic Data in City Planning Departments: The Volume and Use Related to Advancements in Geographic Information Systems (GIS) Technology" (1992). *Dissertations and Theses*. Paper 1352.

[10.15760/etd.1351](https://pdxscholar.library.pdx.edu/open_access_etds/10.15760/etd.1351)

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact [pdxscholar@pdx.edu](mailto:pdxscholar@pdx.edu).

GEOGRAPHIC DATA IN CITY PLANNING DEPARTMENTS: THE VOLUME AND  
USE RELATED TO ADVANCEMENTS IN GEOGRAPHIC INFORMATION  
SYSTEMS (GIS) TECHNOLOGY

by

FAWZI SAEED KUBBARA

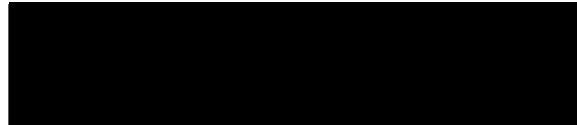
A dissertation submitted in partial fulfillment of the  
requirements for the degree of

DOCTOR OF PHILOSOPHY  
in  
URBAN STUDIES

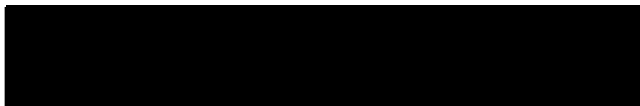
Portland State University  
1992

TO THE OFFICE OF GRADUATE STUDIES:

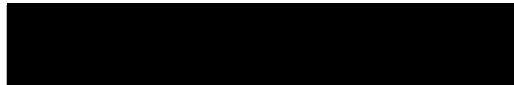
The members of the Committee approve the dissertation of Fawzi Saeed Kubbara presented March 6, 1992.



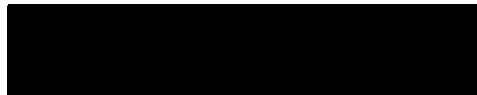
Kenneth J. Dueker, Chair



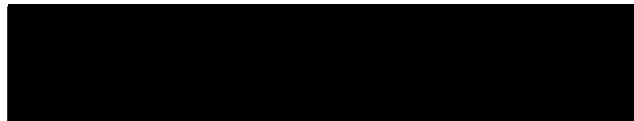
William A. Rabiega



Deborah A. Howe



D. Richard Lycan



Robert C. Liebman

APPROVED:



Nohad A. Toulan, Dean, School of Urban and Public Affairs



C. William Savery, Vice Provost for Graduate Studies and Research

AN ABSTRACT OF THE DISSERTATION OF Fawzi Saeed Kubbara for the Doctor of  
Philosophy in Urban Studies presented March 6, 1992.

Title: Geographic Data in City Planning Departments: The Volume and Use Related to  
Advancements in Geographic Information Systems (GIS) Technology.

APPROVED BY THE MEMBERS OF THE DISSERTATION COMMITTEE:

[REDACTED]

Kenneth J. Dueker, Chair

[REDACTED]

William A. Rabiega

[REDACTED]

Deborah A. Howe

[REDACTED]

D. Richard Lycan

[REDACTED]

Robert C. Liebman

Many local planning departments have acquired and put into use advanced  
automated geocoding and Geographic Information Systems (GIS) to store, process, map,  
and analyze geographic data. GIS technological advancements in hardware, software,

and geographic databases – specifically, in geocoding methods to reference street address data to geographic locations – enable data to be integrated, mapped, and analyzed more efficiently and effectively. Also, technological advancements depend on organizational and institutional environments. The relationships between technological advancements and technical (data mapping and analysis), organizational, and institutional environments are not clear. The purpose of this study is to explain these relationships to help planning and development directors make better decisions in acquiring and using advanced geocoding and GIS technology. The findings are based on a mail survey of planning and development departments in cities with populations of 50,000 or more in the United States.

The study found that planning departments with advanced geocoding and GIS technology are capable of conducting advanced geocoding applications. Data can be tabulated, aggregated, linked, and modeled for mapping and planning. Geocoding to aggregate data to small geographic areas helps by providing required and up-to-date information to solve urban problems. However, the study did not find that advanced geocoding systems enhance data quality as measured by spatial resolution and volume. Further studies are needed to explore this issue.

The adoption and implementation of advanced geocoding and GIS technology are influenced by organizational and institutional environments. Large cities have more experience with hardware, software programs, computer professionals, and training programs, but they are dependent on centralized systems from an earlier computer era. Consequently, more recent entrants to using computers for geographic data processing are emerging rapidly. As technology is becoming more advanced, hardware and software costs are declining. Some of the organizational and institutional issues are eliminated while new ones are emerging. As a result, small area cities are adopting advanced

geocoding and GIS technology more rapidly than they were previously, and sometimes they surpass large cities.

This study improves understanding of automated street address geocoding methods and how these methods are related to advancements in GIS technology. The study also examines how technical, organizational, and institutional environments are interrelated in adopting and using geocoding and GIS technology. The challenge in the 1990s will not be how to fund and acquire a GIS, but how to integrate all of the pieces in order to make the technology work properly.

## DEDICATION

In the name of Allah “God”

Most Gracious, Most Merciful

I dedicate this dissertation to the memory of my late mother, for her love and encouragement during my first three years of study at Portland State University.

May Allah “God” bless her soul.

## ACKNOWLEDGMENTS

It is an honor for me to express my gratitude to my father Saeed, to my brothers Jameel and Adil, and to my sisters. To all of them, I offer my thanks for the intellectual, emotional, spiritual, and financial support they have provided me. Without them, I would not have had the opportunity to reach my lifelong goal.

I would like to express my thanks and respect to my adviser and the members of my dissertation committee. Special thanks go to Professor Kenneth J. Dueker for his advice and guidance in conceptualizing and formulating my research. His knowledge and his patient assistance with selecting a topic, studying the literature of geocoding and GIS technology, facilitating questionnaire design and implementation, and analysis and presentation of the findings are very much appreciated. Thanks go to the committee members for their comments and suggestions, which made this research more valuable and durable. Also, I would like to express my appreciation to the School of Urban and Public Affairs for its financial assistance, and to the Center for Urban Studies for supporting my survey expenses.

Finally, I would like to thank my wife for her love, and my two sons Aahd and Safwan for patiently enduring both parents' dissertations.



## TABLE OF CONTENTS

	PAGE
ACKNOWLEDGMENTS.....	iii
LIST OF TABLES.....	vii
CHAPTER	
I INTRODUCTION.....	1
II REVIEW OF RELATED LITERATURE .....	5
Geocoding: Theoretical Literature.....	6
Geocoding in the 1960s and 1970s .....	8
Geocoding in the 1980s and 1990s .....	9
Geographic Information Systems (GIS).....	10
Geocoding: Data Volume and Use.....	13
Data Volume .....	14
Data Use.....	16
Geocoding: Empirical Research.....	18
Literature Summary and Research Hypotheses .....	22
III RESEARCH METHODS.....	26
Study Population.....	26
Primary Data Collection.....	27
Survey Design.....	28
Survey Implementation.....	29
Survey Response.....	30
Secondary Data Collection.....	31
Data Analysis .....	31

IV	DESCRIPTIVE RESULTS.....	32
	Technical Environment.....	33
	Types of Departments.....	33
	Data Volume and Use.....	36
	Hardware Configurations.....	40
	Application Software.....	42
	Geographic Base File (GBF).....	44
	Digital Map Base Used for Display.....	45
	Organizational Environment.....	45
	City Organization.....	46
	Locations of GIS in the Metropolitan Area.....	48
	Geographic Base File Maintenance.....	49
	Institutional Environment.....	50
	Full-Time Employees.....	50
	Types of Training.....	51
	Planning Department Grouping.....	53
	Summary.....	60
	Conclusions.....	62
V	ANALYTICAL RESULTS.....	65
	Variable Operationalizing and Data Encoding.....	65
	Ordinal Variables.....	66
	Ratio Variables.....	69
	Nonparametric Statistics.....	71
	Kendall's Tau-b.....	73
	Stuart's Tau-c.....	73
	Goodman and Kruskal's Gamma.....	73
	Spearman's Rho.....	74
	Hypothesis Testing.....	75
	Methods of Geocoding and Data Use.....	76
	Methods of Geocoding and Spatial Resolution.....	77
	Methods of Geocoding and Data Volume.....	78
	Spatial Resolution and Data Use.....	79
	Spatial Resolution and Data Volume.....	80

Methods of Geocoding and City Area.....	81
Methods of Geocoding and City Population.....	82
Methods of Geocoding and Department Budget.....	83
Summary .....	83
Conclusions.....	84
<b>VI DISCUSSION AND CONCLUSIONS.....</b>	<b>86</b>
Theoretical Implications.....	87
Technical Issues .....	88
Organizational and Institutional Issues.....	90
Implications for Development and Acquisition.....	91
Technical Issues .....	92
Organizational and Institutional Issues.....	94
Research Needs.....	96
Technical Research .....	96
Organizational and Institutional Research.....	97
Conclusions.....	98
<b>SELECTED BIBLIOGRAPHY .....</b>	<b>99</b>
<b>APPENDICES</b>	
A SURVEY INSTRUMENT AND COVER LETTERS.....	108
B SUMMARY OF STATISTICS.....	119
C CITIES SURVEYED.....	121

## LIST OF TABLES

TABLE	PAGE
I     Types and Sources of Data .....	14
II    Mailing Schedule for Surveys.....	29
III   Completed Surveys Returned each Week.....	30
IV    Percentages of Manual Departments Using Manual Geocoding Methods.....	34
V     Percentages of Automated Departments Using Automated Geocoding Methods.....	35
VI    Percentages of Spatial Resolution by Automated Departments .....	37
VII   Percentages of Data Types by Automated Departments.....	38
VIII   Types of Geocoding Applications by Automated and Manual Departments.....	38
IX    Types of Census Applications by Automated and Manual Departments.....	39
X     Types of Hardware Configurations by Automated and Manual Departments.....	40
XI    Average Number of Years Since Each Type of Hardware Configuration Was Introduced by Automated and Manual Departments.....	41
XII   Average Number of Seats for Each Type of Hardware Configuration by Automated and Manual Departments .....	42

XIII	Types of Application Software by Automated and Manual Departments.....	43
XIV	Average Number of Years Since Each Type of Application Software Was Introduced by Automated and Manual Departments.....	43
XV	Types of Geographic Base Files by Automated Departments.....	44
XVI	Types of Digital Map Base Used for Displaying Data by Automated Departments.....	45
XVII	Centralized and Decentralized GIS Units by Automated and Manual Departments.....	46
XVIII	City Departments with Centralized and Decentralized GIS Units by Automated and Manual Departments .....	47
XIX	Locations of GIS in Other Organizations in the Metropolitan Area by Automated and Manual Departments .....	48
XX	Maintenance Locations of Geographic Base Files by Automated Departments.....	49
XXI	Types of Full-Time Employees by Automated and Manual Departments.....	50
XXII	Average Number of Each Type of Full-Time Employees by Automated and Manual Departments .....	51
XXIII	Types of Training by Automated and Manual Departments .....	52
XXIV	Average Number of Hours for Each Type of Training by Automated and Manual Departments.....	52
XXV	Definition of Variables Used in K-Means Cluster Analysis.....	54
XXVI	Cluster Analysis Results for Planning Departments.....	55

XXVII	Methods of Geocoding and Their Rankings for Automated and Manual Departments.....	67
XXVIII	Data Uses and Their Rankings for Automated and Manual Departments.....	67
XXIX	Spatial Resolutions and Their Rankings for Automated Departments.....	69
XXX	Data Volume Ranking for Automated Departments.....	70
XXXI	City Area, City Population, and Department Budget Rankings for Automated and Manual Departments .....	71
XXXII	The Relationship Between Geocoding Methods and Data Use for Automated and Manual Departments .....	77
XXXIII	The Relationship Between Methods of Geocoding and Spatial Resolution for Automated Departments .....	78
XXXIV	The Relationship Between Methods of Geocoding and Volume of Data for Automated Departments.....	79
XXXV	The Relationship Between Spatial Resolution and Data Use for Automated Departments.....	79
XXXVI	The Relationship Between Spatial Resolution and Data Volume for Automated Departments.....	80
XXXVII	The Relationship Between Methods of Geocoding and City Area for Automated and Manual Departments .....	81
XXXVIII	The Relationship Between Geocoding Methods and City Population for Automated and Manual Departments.....	82
XXXIX	The Relationship Between Geocoding Methods and Department Budget for Automated and Manual Departments.....	83

## CHAPTER I

### INTRODUCTION

The demand for valid, timely, and detailed geographic data within local planning departments is increasing rapidly due to the increase in information required for planning, management, and decision making. Also, shrinking revenues increase the pressure on planning and development directors to produce such data from different sources in an efficient and effective manner (U.S. Department of Housing and Urban Development, 1968; Dueker, 1980; Teicholz and Berry, 1983; American Congress of Survey and Mapping [ACSM], 1989).

Geographic data represent geographic features such as points, lines, areas, and grid cells and are linked with attribute data using one or more of the available geocoding (geographic coding) methods. Attribute data such as population, housing, transportation, land use, and many urban data files that include street addresses are linked (geocoded) with these geographic features for tabulation, mapping, and analysis. Street address data and automated street address geocoding systems are the focus of this study.

Automated street address geocoding systems are computerized information systems which assign geographic codes such as census tract numbers, city block numbers, or x,y coordinates to data records using street addresses as location identifiers. Street addresses are the linkages that tie data with geographic locations on maps. These geographic locations may take the form of areas such as census tracts, city blocks, or ownership parcels; the form of segments such as street segments or city block sides; or the form of points such as street intersections or facility locations. For years, data were

geocoded manually. However, since computer systems evolved about four decades ago, the process of geocoding has increasingly been conducted automatically. A geocoding system requires at least three items: a map file, a data file, and a software program to perform the linkage.

In the last few years, geocoding systems have become more advanced. Increasingly, map files not only contain the geometry (x,y coordinates), but are structured topologically, i.e., containing the relations of the points, lines, and areas. Data files are available in digital format, and geocoding programs are available on microcomputers with much greater flexibility and accuracy. These three advancements have led to more efficient and accurate geocoding of street addresses.

Street addresses can be geocoded by matching addresses with geographic locations in a batch or by interactive processing, and in a GIS environment addresses can be geocoded by digitizing, overlaying, or matching them with geographic locations with much greater accuracy and flexibility. Matched addresses with accurate positional locations enable users to tabulate, aggregate, integrate, and model address data for utilization in many planning applications such as mapping and spatial analysis.

As a result, many local governments, specifically local planning departments, have acquired and put into use advanced geocoding and Geographic Information Systems (GIS) to improve the use, quality, and currency of data for land use and transportation planning, facility location, housing and urban development, and many other local planning applications. Many types of up-to-date data referenced by small geographic areas have become available and accessible with the use of GIS and advanced geocoding systems (Dueker, 1980; Simutis, 1980; Cater, 1974).

The relationships between advancements in geocoding systems (methods of geocoding) and geographic data mapping and analysis are not clear. Huxhold (1991),



Kinzy (1977), and Cooke (1980) stated that efficient and effective data volume and use depend on advancements in geocoding methods, while Barb (1974) and Kraemer and Modeleski (1975) stated that advancements in geocoding systems depend on the organizational and institutional environments. For example, advanced geocoding systems and GIS allow street addresses to be geocoded by small geographic areas or unique coordinate locations; this capability increases data volume, enhances spatial resolution, and improves data use. Consequently, detailed and up-to-date data needed by planners, managers, and decision makers can be available and accessible (Cater, 1974; NRC, 1980; Simutis, 1980; Dueker, 1980; Eveland, 1980).

Empirical studies, on the other hand, show that technological advancements in geocoding are a product of departments' organizational and institutional environments. Cities and counties with large populations, large annual operating budgets, and large numbers of full-time employees use computer and geocoding technology more frequently than do smaller cities (U.S. Department of Housing and Urban Development, 1971; Kraemer and Modeleski, 1975).

Clearly, research is needed to examine the relationships between technological advancements in geocoding systems and technical (geographic data mapping and analysis), organizational, and institutional environments. A study of these relationships will help local officials make better decisions in acquiring and developing geocoding and GIS technology needed for local planning applications such as land use, housing, urban renewal and development, and many other planning applications. The current analysis surveyed 427 planning departments in cities with populations of 50,000 or more in the United States.

This dissertation is organized into six chapters. Chapter II lays the foundation for the study by reviewing theories and empirical research on geocoding and GIS technology.

The emphasis is on automated street address geocoding systems. A summary of the literature and a discussion of the research hypotheses are also included in this chapter. Chapter III presents the study population and discusses the collection of primary data including survey design, survey implementation, and survey response; secondary data collection; and methods of analysis. Chapter IV identifies and compares planning departments in terms of technical, organizational, and institutional environments. Chapter V analyzes the relationships between geocoding methods and data use, spatial resolution, and data volume; between spatial resolution and data use and volume; and between geocoding methods and area, population, and department budget. Chapter VI discusses the implications of the study for theories of geocoding, for GIS development and acquisition, and for future research.

## CHAPTER II

### REVIEW OF RELATED LITERATURE

The automated street address geocoding literature can be divided into two periods of time. The first, from the 1960s to the 1970s, describes methods of geocoding in mainframe computers for general mapping and geocoding applications. The second, from the 1980s to the early 1990s, describes street address geocoding methods within the context of GIS for advanced mapping and planning applications. Current geocoding literature is referred to as GIS literature.

Reviewing the literature on automated street address geocoding methods was not easy, due to three problems. First, literature on geocoding was available in at least three different areas, including information systems, geography, and urban planning. Second, literature on geocoding was written and documented by three groups of people, namely academicians, vendors, and users. Third, the multidisciplinary nature of the field caused a scarcity in the amount and types of empirical and theoretical research, since few professionals are conducting research in this field. Very few theoretical and empirical studies related to this study were found. Moreover, none of the previous research studied data detail, types, and use as a result of advancements in GIS technology. This was probably due to the infancy of this technology, which became widely used only in the last few years. The results offered here are the first set of findings on what will become an important area of research.

The chapter is divided into three sections. In the first section, theoretical literature on geocoding is reviewed, including literature on the U.S. Bureau of the Census

Geographic Base Files (GBFs), GIS, and mapping systems. The purpose of this section is to define each type of geocoding method and describe how advancements have occurred. The second section introduces data volume and use related to geocoding. The third section reviews empirical research on geocoding and GIS in local governments. Finally, hypotheses are derived and included in the last section.

The review in the first section was derived primarily from two well-known documents written on geocoding in the 1970s: Charles Barb's dissertation "Automated Street Address Geocoding Systems: Their Local Adaptation and Institutionalization" in 1974, and Stephen Kinzy's report "Geoprocessing in Local Governments: A Feasibility Study for the City of Omaha" in 1977. Geocoding literature of the 1980s and early 1990s was derived from GIS literature available from conferences, private and public reports, vendor publications, and more recently, GIS books.

### GEOCODING: THEORETICAL LITERATURE

A geoprocessing (geographic analysis and data processing) system is an information system that can store, manipulate, and display planning and map data with analytical and retrieval capability for land use and transportation planning, statistical mapping, decennial census tabulation, and other geographic analysis (Kinzy, 1977). Horwood (1967) defined the components of such systems, when they are applied to planning, as:

- 1) A geocoding system to convert data input by street address location identifiers to geographic coordinates that allows easy retrieval of information.
- 2) A query system to manipulate large databases.

- 3) An automated graphic display system for map making and data display.
- 4) A plan test system to test alternative proposals regarding employment distribution, residential densities, transportation facilities, and growth.
- 5) A planning operation system for everyday needs, document retrieval, report generation, and statistical reports.
- 6) A capital improvement and work-scheduling system to integrate planning and processing of all departments.

The heart of a planning information system is the geocoding system. Geocoding is the procedure of referring data to geographic locations, using street addresses as the location identifiers. The procedure ranges from relatively basic (such as referencing area codes to data records in a batch or interactive process) to complex (such as assigning x,y coordinates using the U.S. Bureau of the Census GBFs that have been developed in the last three decades) to extremely complex (such as geocoding within a GIS context using segment- and parcel-based map bases) (Eveland, 1980).

Geocoding of data by street addresses can be conducted manually or automatically. Manual geocoding is typically done directly by recording data with locations on maps, or indirectly by recording geographic codes in data records (Kinzy, 1977). Automated geocoding is the automation of the manual process of assigning data to geographic entities using their nominal or positional geocodes. A nominal geocode could be a census tract number or a street name. A positional geocode, on the other hand, could be a location of the x,y coordinates in the state plan coordinate system (Barb, 1974; Kinzy, 1977). Kinzy (1977) and Cooke (1980) contended that positional geocodes afford broader utility and greater flexibility in data processing than do nominal geocodes.

The automated assignment of positional geocodes on data records is the most common method used among local governments. This method requires the following:

- 1) a map file such as the U. S. Bureau of the Census Geographic Base File (GBF),
- 2) data files such as employment files, building permit files, etc., and 3) software programs such as ADMATCH and UNIMATCH. The availability of the U.S. Bureau of the Census Geographic Base Files and their associated geocoding software programs encouraged local departments to automate the geocoding process that was previously done manually (Eveland, 1980; Clark, 1967). GBFs and geocoding programs have been developed, advanced, and used by many local and regional planning departments for planning, mapping, spatial analysis, and data integration.

#### Geocoding in the 1960s and 1970s

The GBF files, such as the Address Coding Guide (ACG) and the Dual Independent Map Encoding (DIME), were designed and developed by the U.S. Bureau of the Census in the 1960s and 1970s. The design of the ACG started more than three decades ago. It consisted of block side records between two intersections for all streets within the urban area (Cooke, 1980). The ACG file did not use the concept of topology or an accurate geodetic system. Consequently, blocks were not plotted accurately and spatial analysis was impossible. To overcome these problems, the Census Bureau designed a new system called Dual Independent Map Encoding (DIME). This second-generation approach describes the urban street network in terms of its main components: segments, nodes, and blocks. It describes the segments of the network in topological and topographical terms; acts as a reference file among coordinates, postal addresses, and block numbers; and serves as an appropriate cartographic base for mapping local street map files.

The DIME file was designed to geocode street addresses for the 1980 census and to facilitate linking of data files through the use of spatial reference files. The problem associated with the DIME file was basically inefficiency due to the system's complexity, slow system response, and large number of positional errors. Another problem noted was the lack of match between user problems and the system configuration (Kinzy, 1977; Eveland, 1980; Aangeenburg and Bomberger, 1975).

#### Geocoding in the 1980s and 1990s

In the 1980s, the Census Bureau designed a new file to automate mapping and related geographic activities. The new file is known as Topologically Integrated Geographic Encoding and Referencing (TIGER). It overcomes previous file problems because TIGER files are relationally structured (to encode geographic entities and reference them to one another), and the use of the U.S. Geological Survey (USGS) 1:100,000-scale maps meets the national map accuracy standards (Marx, 1986b; USGS, 1985). These two features of the TIGER file reduce the problem of positional accuracy of the previous GBFs. Thus, street addresses have more accurate positional locations; this accuracy is important for planning data geocoded by small areas.

The most sophisticated geocoding systems of the 1980s and 1990s are those that process the U.S. Bureau of the Census GBF (e.g., TIGER) with one or more of the following mapping systems: Thematic Mapping, Computer-Aided Design and Drafting (CADD), Land Information Systems (LIS), and Multipurpose Cadastre Systems in a GIS environment. The GIS capabilities of creating, editing, deleting, linking, overlaying, and displaying geographic data available in the above mapping systems simultaneously increase geocoding efficiency and effectiveness. Also, higher positional and relative accuracy of the above mapping systems provides accurate positional locations for street

addresses, thereby enabling planning data mapping, analysis, and integration to be more accurate. A brief definition of a GIS, and how each of the mapping systems is related to geocoding in the GIS context, are provided in the following section.

### Geographic Information Systems (GIS)

A GIS allows users not only to display geographic and non-geographic data, but also to perform a number of technical operations on the data such as creating, editing, and deleting geographically structured data; linking locational and attribute data for mapping and spatial analysis including map overlay analysis; and displaying and drawing geographic information (Dueker and Kjerne, 1989).

To facilitate understanding of how a GIS system is used to implement advanced geocoding, the components of GIS systems, as defined by Teicholz and Berry (1983) and Star and Estes (1990), are listed below.

- 1) Data acquisition systems responsible for graphic and statistical data, encoding, and input processing,
- 2) Data management systems responsible for report generation, security, data integration, and statistical report functions,
- 3) Data manipulation and analysis systems responsible for projections, transformations, combinations of different types of data such as polygon overlay, and statistical analysis, and
- 4) Data display systems responsible for output maps, graphs, and tabulated information on a variety of output media.

Because of the above GIS components and technical operations, street addresses can be geocoded in two different ways. The first is digitizing, in which each street address is digitized to a location using a digitizing device. The digitizing process is



available in data acquisition systems, in which digitizing is one way to enter geographic data and create map bases called coverages. The second is overlaying, in which a map coverage with street addresses is overlaid on another map coverage. The overlaying process is available in data manipulation systems, in which street addresses are transferred from one coverage to another and new attribute data are assigned to street address records. Both coverages have to be geodetically referenced, and street addresses have to have accurate locations so that they can be transferred to the right locations in the new coverages.

The power of GIS is its ability to process many different types of maps which help to transfer street addresses from one map system to another so that more attribute data can be related to the same geographic locations. Following are definitions of some maps used in GIS for geocoding.

Thematic Mapping. These types of maps are designed for analysis and graphical display of geographic data, aggregated to geographical or postal boundaries such as counties, census tracts, and ZIP codes, or customer regions such as territories, tax parcels, and school districts (Cook, 1989). These systems are used to display areal data, and are not capable of performing locational or spatial analysis, nor aggregating or disaggregating data.

Computer-Aided Design and Drafting (CADD). CADD maps are basically drawing tools. Their databases contain only cartographic information and are used for general-purpose map management with Automated Mapping and Facility Management (AM/FM) systems (Dangermond, 1989). The advantage of CADD systems is the accurate registration of the cartographic data. There are many local planning departments using CADD map bases for data registration with high positional accuracy. CADD data can be layered geographically for presentation, but cannot be related logically.

Land Information Systems (LIS). LIS systems are concerned with land records that include resources, land use, environmental impact, and fiscal data. Attribute and geographic data are interrelated by location identifiers. If cadastral and property information are included, the LIS may be referred to as Multipurpose Cadastre Systems (Dueker and Kjerne, 1989; Moyer, 1980; Huxhold, 1991).

Multipurpose Cadastre Systems. The concept of Multipurpose Cadastre Systems is based on the processing of parcel maps (HUD, 1968; NRC, 1980). The Multipurpose Cadastre Systems consist of the following components:

- 1) The reference frame that consists of a system of monuments having geodetically-derived coordinates and that permits defining the spatial location of all land-related data.
- 2) The large-scale mapping series, consisting of a family of planimetric and topographic maps at scales from 1:500 to 1:25,000 to permit the graphical representation of the land-related data.
- 3) The cadastre overlay, consisting of a socialized series of maps delineating the current state of property ownership. The individual building for the overlay is the cadastre parcel. A unique identification number is associated with each cadastre parcel.
- 4) The unique parcel identification number that provides a key for linking each cadastre parcel to various land data files or registers. These records may contain information about land ownership, use, cover, assessment, and other such attributes.
- 5) A series of registers or land data files, each including a parcel index for information retrieval and linkage with information in other data files (McLaughlin, 1975, cited in NRC, 1980).

The key principle underlying the concept of a multipurpose cadastre is that the cadastre layer contains geographic and attribute information associated with each parcel. Attribute data, such as information about population, housing, travel, land use, and many other data files, can be associated with each parcel. This is possible when segment- and parcel-based systems are overlaid in a GIS (Bryan et al., 1991; Barnell, 1989; Somers and Eichelberger, 1987). For example, street addresses available in segment-based systems can be assigned to each ownership parcel. Attribute data in both files can be integrated for each parcel; this integration makes data mapping and analysis more accurate, reliable, and efficient. Also, data can be processed, mapped, and analyzed based on their unique coordinate locations; these capabilities provide accurate and detailed information to planners and managers who have not had it before.

The advances in geocoding systems presented in this section should increase data volume and use. Data about population, housing, transportation, and land use can be geocoded with small geographic areas or unique x,y coordinates needed for local and regional planning. Also, data geocoded by small geographic areas can be used to integrate data from different sources. Following are definitions of data volume and use. Data volume is defined in terms of spatial resolution and types of data, and data use is defined in terms of types of geocoding applications.

### GEOCODING: DATA VOLUME AND USE

The major purpose of this study is to examine whether advancements in geocoding methods improve data volume and use. The previous section introduced and explained the components and advancements of geocoding systems. This section defines and illustrates how data volume and use are related to advancements in geocoding methods.

### Data Volume

Data volume is related to the types of data and their spatial resolutions. Data types include population, housing, physical features, facilities and services, and resources. They are collected and captured by many departments and agencies at the local, state, and federal levels (see Table I). Each data record has a street address and is referenced primarily to geographic areas for mapping and analysis. Information can be made available by geographic area for planning, management, and decision making. The types and sizes of geographic areas are usually determined by the needs and types of applications. Addresses are often referenced by segments and points for transportation and facility planning; however, this study focused on areal geography.

TABLE I  
TYPES AND SOURCES OF DATA

<u>Types</u>	<u>Data</u>	<u>Sources</u>
Population	Age, family structure, vital statistics, income, level of education	U.S. Census, state and local offices, health departments
Housing	Type, cost, zoning, quality, building permits	Assessor's office, local planning agencies
Community physical features	Land use, utility, environmental quality, transportation	Planning agencies, highway departments
Community facilities and services	Health care, recreation, welfare support, volunteer organizations	Hospital records, parks departments, federal agencies' reports, local surveys
Community resources	Public safety personnel, funding for community facilities, property tax base, commercial revenues	Local budget office, federal agencies

Source: Simutis, 1980

Spatial resolution is defined in terms of the sizes of geographic areas such as census tracts, city blocks, and ownership parcels (Becker, 1980; Dueker, 1980). The more types of data that are related to the same geographic area in the database, and the

smaller the geographic area, the larger the volume of data will be. The following explains how data types are interrelated with spatial resolution as explained by Kinzy (1980) and Dueker (1988).

Level I Spatial Resolution. At this level, areas such as census tracts, traffic zones, and neighborhoods are used for geocoding. Positional accuracy of 100 to 1000 feet and relative accuracy of 10 to 100 feet are required. Data registration does not rely on coordinates. Nominal geographic codes such as census tract numbers, traffic zones, or ZIP-code numbers are used to spatially register geocoded data. Data geocoded to these areas are used for planning applications such as transportation, land use planning, and marketing (Dueker, 1988).

Level II Spatial Resolution. At this level, city blocks and street segments are used for geocoding. Positional accuracy of 10 to 50 feet and relative accuracy of 1 to 10 feet can be obtained. Geocoding consists of converting street address-referenced data to nominal geographic codes, such as city block numbers or segment numbers, or to approximate coordinates based on interpolating between intersections using address ranges. Data geocoded at this level with this accuracy are used for routing and locational analysis. Data volume is greater at Level II than at Level I due to the smaller geographic areas (Kinzy, 1980; Dueker, 1988). Data at this level could be used for tabulation, mapping, and spatial analysis, but mapping and spatial analysis may not be accurate due to the low positional and relative accuracy involved at this level of resolution.

Level III Spatial Resolution. Three categories of resolutions are involved at this level: Facility Management, of concern to organizations that maintain and manage facilities; Tax Mapping, of concern to assessors and others who maintain property ownership records; and Engineering, of concern to public works organizations that deal with maintaining segment-based maps for use in engineering. Positional accuracy of 0.1

to 10 feet is required, and coordinate systems are used for spatial registration of data. The volume of data is greater than at Levels I and II due to smaller geographic areas and units of observation. Also, the cost of Level III database design is higher than that of the first two levels due to the required registration accuracy (Kinzy, 1980; Dueker, 1988).

Level III data are used for tabulation, mapping, and spatial analysis. Also, the data can be aggregated or disaggregated easily and precisely due to the high positional and relative accuracy required at this level of resolution. For example, data geocoded by street addresses using accurate x,y coordinates are assigned to data records, thereby allowing data to be located at the exact locations. The accurate locations of data increase flexibility in data integration with other data when they are geocoded at the same locations.

Three levels of spatial resolution were introduced and explained above. Each level of resolution geocodes data differently. At the first level, nominal geocodes are utilized for geocoding, while the other two levels utilize positional geocodes with or without accurate coordinate registrations for geocoding. Data volume increases as resolutions and types increase.

### Data Use

Geocoding facilitates data aggregation, mapping, analysis, and display needed for planning applications. Planning data can be aggregated by area, merged with other data files, analyzed statistically or graphically, and displayed on screens or drawn by plotters (Barb, 1974; Kinzy, 1977).

For example, planning departments geocode employment, auto registration, travel origin and destination, and building permit data to describe the type, number, and location of employees in a region, to describe the residential location of the vehicle owner in a region, to describe the general travel pattern within a region, and to describe building

permit locations. Other organizations such as hospitals, school districts, and private agencies geocode data by census tracts or point coordinates to define service area boundaries and facility locations and to tabulate data for school location and service boundary analysis (U.S. Bureau of the Census, 1971c).

Geocoding within a GIS context enables new and advanced planning applications. For example, the availability of low-cost TIGER/Line files, which are extracts of selected information from TIGER files, and of software such as ARC/INFO and SAFARI, are believed to have increased planning applications as well as data uses (Simmons, 1989; Teicholz and Berry, 1983). Another example is Tacoma's Geobased Systems (GBS), which automate parcel-level ownership data to provide information to all departments in the city. With GBS, applications such as parcel ownership inventory, spatial analysis, facility management, and routing can be performed (Wood, 1989).

In short, the uses of data as a result of street address geocoding can be grouped into at least four levels, from the least advanced to the most advanced, as follows:

- 1) Data can be tabulated by statistical areas without the results being mapped.  
For example, addresses can be matched with traffic zones to update employee data, or addresses can be tabulated for mailing to neighborhoods.
- 2) Data can be aggregated to statistical areas and mapped. For example, addresses can be assigned to areal units such as census tracts, ZIP codes, etc., for summary and mapping.
- 3) Data can be matched to coordinates and/or network nodes. For example, addresses can be assigned to points and mapped as incidents.
- 4) Data can be linked with coordinate-based systems for modeling and other application programs. For example, data can be geocoded to point locations and used as inputs to other analytical or modeling processes.

Advanced geocoding and GIS technology provides data for tabulation, mapping, and spatial analysis. Also, these data, with accurate registration, can be aggregated by small areas and integrated with other data from different sources needed for advanced planning applications.

### GEOCODING: EMPIRICAL RESEARCH

The use of computer systems in local governments started in the 1950s and 1960s. Early studies focused on estimating and evaluating the automation of the manual process for routine operations in administration and finance departments. For example, with greater speed and accuracy in processing huge amounts of data for calculation and sorting, local governments adopted these technologies and thereby increased their performance and productivity (Kraemer and King, 1977b). Current studies are being conducted on geocoding systems in local planning departments at the national and local governmental levels, but for the purpose of adoption and implementation; these studies are reviewed in this section.

In the 1970s, research was undertaken to examine the impact of the increased use of information systems in local governments. Two studies at the national level were conducted. The first study was funded through an Urban Information System Inter-Agency Committee (USAC) contract awarded by the Department of Housing and Urban Development. This study conducted interviews in seventy-nine cities which had experience with information systems (U.S. Department of Housing and Urban Development, 1971).

Cities were ranked based on factors related to the hardware and software, data management, municipal support, and environments of municipal Data Processing (DP) departments. It was found that the highly ranked cities had large populations, large



annual operating budgets, large numbers of full-time employees, and strong executive forms of government (e.g., city managers); were located in the south central, western, and northwestern regions of the U.S.; developed in-house DP and training; had large DP expenditures and staffs; and consolidated all DP activities into a single unit. However, larger DP expenditures did not necessarily produce a greater number of applications.

Computer applications were classified into four groups and were given functional titles as follows: Public Finance and Administration, Physical and Economic Development, Human Resources Development, and Public Safety. The study indicated that the Public Finance and Administration and Public Safety functions were the most computerized and had many applications because these were the areas where the computers were developed. The area of Physical and Economic Development rated third, with computerized applications including planning, zoning, subdivision regulations, building inspections, engineering, park and street maintenance, and utility operations (U.S. Department of Housing and Urban Development, 1971).

The second study to evaluate the use and implementation of geocoding systems was carried out by the Public Policy Research Organization, a research center at the University of California, Irvine (Kraemer and Modeleski, 1975). This project, known by its acronym, URBIS (for Urban Information Systems), surveyed existing geocoding systems in U.S. cities and counties, with funding from the Research Applied to National Needs (RANN) Division of the National Science Foundation (NSF).

The URBIS survey was sent to all U.S. cities with populations of more than 50,000 and all counties with populations greater than 100,000 (713 local governments). A total of 477 cities and counties responded to the applications survey (67%). In the survey questionnaire, four geocoding systems were surveyed: 1) Street Address

Conversion System (SACS), 2) Address Coding Guide (ACG), 3) Dual Independent Map Encoding (DIME), and 4) Address Matching (ADMATCH). The most important findings were:

- 1) Use of geocoding systems was strongly related to population size. For both cities and counties, the proportion of geositing (sites with one or more operating geocoding applications) increased as the size of each population increased.
- 2) Geosites had larger budgets, larger per-capita budgets, larger DP budgets, and larger DP staffs than average.
- 3) The number of geosites was relatively small. Of the total 477 cities and counties responding (67% return), only 83 (17%) reported one or more geocoding applications. Cities showed a slightly higher frequency of geocoding use than counties.

At the local level, Barb (1974) conducted a case study of two cities using geocoding systems, Seattle and Philadelphia. He found that 1) local implementation of geocoding technology is progressing slowly; 2) existing local agencies' application of geocoding is elementary; 3) the trend in existing local systems' evaluation is characteristically uncoordinated, isolated, and dedicated to the limited internal objectives of local agencies; 4) systems are highly complex; and 5) local agencies are unprepared to fully develop geocoding systems technology. In short, the problem of adopting this technology in local governments is not technical, but organizational and institutional.

Barb's second major finding was that geocoding technology was available for adoption and implementation, but only elementary utilization was implemented -- data aggregation by areas or polygons and merging files, but not computer mapping and geographic analysis. He recommended the following: 1) system capitalization in which

funding must be stable, long-term, and local; 2) system identity in which staff must be independent from other departments' staffs; 3) policy formulation in which technical and management objectives are made clear; 4) a system service role in which technical and management aids are available; and 5) system ownership in which participants have access, usage, design, and operation as recognition. In short, more management and organizational funding support should be considered before a geocoding system is adopted (Barb, 1974).

In a second study, Barb conducted a survey of eleven cities. The results of this survey showed that only half of the cities participated in the Census Bureau's CUE Update Program, two of the cities used their own geographic reference systems, while the others used geocoding application packages without DIME files as a base (Barb, 1976; cited in Kinzy, 1977).

In the 1980s, GIS became a term commonly used in the literature for information and geocoding systems. The American City and County Annual Survey provides current information on the use of computer technology and GIS systems in particular for all U.S. cities and counties. The 1989 survey concluded that GIS system use had expanded in cities and counties because of GIS's low cost and increasing power and storage capabilities. One-third of the U.S. cities with populations over 75,000 had implemented GIS systems. Of the two-thirds not using GIS systems, seventy-six percent indicated they would do so in the next three years. The study identified institutional barriers to adoption, such as training and knowledge of GIS technology. According to the survey, computers are not desired by city and county officials if no one knows how to use them and if training is a problem (Darnell and Carlile, 1989). The use of GIS systems will continue to increase in local governments if the above conditions change.

A study titled "California Planning Agency Experience with Automated Mapping and Geographic Information Systems" (French et al., 1989) showed that in the last few years, the percentage of computer mapping users had increased from 11.8% in 1982 to 58.8% in 1987-1989. The findings of this survey described the increased use of microcomputers and the increased variety of GIS programs and applications in these agencies. For example, the predominance of applications by the planning agencies were for land parcel mapping (61.1%) and related activities such as permit tracking and vacant land inventorying. The second major use was in the area of traditional planning activities such as general plan preparation, zoning review, and growth monitoring.

#### LITERATURE SUMMARY AND RESEARCH HYPOTHESES

Automated street address geocoding is the process of linking street addresses available in data records with geographic locations on maps by assigning nominal and/or positional geocodes for the purpose of tabulation, mapping, and spatial analysis automatically. Yet the previous review points out that little is known about geocoding within the GIS context. It seems clear that large organizations adopt and use geocoding technology. What is not known is how advancements in geocoding methods are related to technical, organizational, and institutional environments. The literature review suggests several possibilities:

- 1) Geocoding within a GIS context provides broader utility and greater flexibility in data processing (Kinzy, 1977; Cooke, 1980). Data can be tabulated, aggregated, matched with coordinates, and linked to models and other applications. Advanced planning applications such as mapping and spatial analysis can be conducted easily, and investigation is needed.

- 2) Geocoding within a GIS context may improve spatial resolution and increase data volume. Data from different sources could be linked with city blocks and ownership parcels accurately. These relationships have not been investigated.
- 3) High spatial resolution may improve data use. Data geocoded by small geographic areas could be tabulated, aggregated, matched to coordinates, and linked to modeling and other application programs. This view was proposed by Kinzy (1980), Dueker (1988), Wood (1989), and investigation is needed.
- 4) High spatial resolution may increase data volume. Data volume is greater at Level III spatial resolution than at Levels I and II due to the large number of geographic units and number of observations (Kinzy, 1980; Dueker, 1989). This relationship has not been investigated.
- 5) Jurisdictions with large populations, large budgets, large numbers of full-time employees, and many other organizational and institutional characteristics were considered as factors in adopting geocoding technology at the local government level. Some of the empirical research supports this view (Kraemer and Modeleski, 1975) and an update is needed, especially for GIS.

The following hypotheses are derived from the above conclusions for investigation and analysis.

Hypothesis 1. Methods of geocoding are significantly correlated with data use. Planning departments using GIS are better able to conduct advanced geocoding applications such as matching addresses to coordinates and aggregating address-based data for modeling.

Hypothesis 2. Methods of geocoding are significantly correlated with spatial resolution. Planning departments with advanced geocoding systems use small geographic areas for mapping and analysis, whereas planning departments with less-advanced geocoding systems can only tabulate addresses by statistical areas without mapping the results, or aggregate data by statistical areas while also mapping the results.

Hypothesis 3. Methods of geocoding are significantly correlated with data volume. Planning departments with advanced geocoding systems can geocode many types of data by small geographic areas for mapping and planning purposes, whereas planning departments with less-sophisticated systems can use only one or two types of data using large geographic areas for specific planning applications.

Hypothesis 4. Spatial resolution is significantly correlated with data use. Data geocoded by small geographic areas can be integrated with other data files and are flexible for modeling for many different planning applications.

Hypothesis 5. Spatial resolution is significantly correlated with data volume. Data geocoded by small geographic areas and associated with a large number of data types increase data volume. Land parcel data can be integrated with other data related to the same location.

Hypothesis 6. Methods of geocoding are significantly correlated with city area size. Large cities are more likely than small cities to acquire and use GIS to store, manipulate, and tabulate large volumes of geographic data.

Hypothesis 7. Methods of geocoding are significantly correlated with city population size. Large cities are more likely than smaller cities to acquire and use advanced geocoding and GIS technologies.

Hypothesis 8. Methods of geocoding are significantly correlated with department budget. Planning departments with large annual budgets can afford advanced GIS technologies more than can departments with small budgets.

This study surveys planning departments in cities with populations of 50,000 or more in the U.S. to examine the relationships between technological advancements in geocoding systems and technical, organizational, and institutional environments.

## CHAPTER III

### RESEARCH METHODS

This study examines the relationships between technological advancements in geocoding methods and data use, spatial resolution, and data volume; between spatial resolution and data use and data volume; and between advancements in geocoding technology and city area, city population, and department budget. The analysis is based on a mail survey of 427 planning and development directors in cities with populations of 50,000 or more in the U.S. Planning and development directors were selected for two reasons: 1) they know which people in their departments are most knowledgeable about geocoding and GIS, and they have the authority to forward the questionnaire to those people, and 2) names of geocoding and GIS managers in city planning departments were not available. Cities with populations of 50,000 or more were selected because they are the most frequent users of geocoding and GIS technology for planning purposes. A survey of planning and development directors should produce a clearer picture of GIS in technical, organizational, and institutional environments.

This chapter presents the study population and discusses the collection of primary data including survey design, survey implementation, and survey responses; secondary data collection; and methods of analysis.

### STUDY POPULATION

The use of information systems in local governments is positively associated with size. Previous studies indicated that cities with large populations tended to use geocoding



systems more than did small cities (Kraemer and Modeleski, 1975). A population of 50,000 was the break point for identifying large and small cities in this study, since this population size was used to identify a Standard Metropolitan Statistical Area (SMSA) by the Bureau of the Census. Therefore, cities with populations of 50,000 or more were selected for this study. Based on the 1990 census, there were 526 cities, townships, and boroughs in the U.S. within this population category. Names and addresses of 427 planning or community development directors were obtained from the Municipal/County Executive Directory 1989/90. Ninety-six names and addresses of mostly towns and boroughs were not available because they were not defined as municipal governments, and three cities were overlooked when the mailing list was compiled. Therefore, the compiled list contained names and street addresses of mainly planning departments in municipal governments available from the above source; these municipal governments were referred to in this study as cities.

Counties and special districts such as schools and regional agencies were not included due to greater size differences, responsibilities, and map scales. For example, city planning departments may use geocoding systems for land use planning, zoning, and subdivisions at a smaller scale than counties, and at a larger scale than special district agencies.

### PRIMARY DATA COLLECTION

The mail survey approach was used because data about geocoding and GIS systems, data volume, and data uses in local governments at the national level were not available. Telephone or on-site interview approaches were not appropriate due to the high cost and the impracticality of site interviews.

One of the major problems associated with the mail survey approach is a low response rate, although Bainbridge (1989) indicated that approximately half of the people who were sent a single mail survey were willing to complete and return it. In order to obtain as many responses as possible, Dillman's (1978) total design method approach was used to increase the response rate. This method is based on a theory of response behavior and an administrative plan to improve the response rate. A well-written questionnaire, careful implementation, and a professional approach are the major criteria for success.

Each planning or community development director was sent a questionnaire with a cover letter explaining the purpose and importance of the study. Each was asked to direct the questionnaire to a person who was familiar with geocoding and GIS systems (see Appendix A for copies of the letters and questionnaire).

### Survey Design

The questionnaire was designed to elicit responses related to technical, organizational, and institutional environments. The questionnaire went through two stages. First, technical questions about types and costs of hardware, software, and database structure were designed to be measured at the interval and ratio levels. The questionnaire was pilot-tested using a sample of six cities that represented different sizes and locations. Two weeks later, follow-up letters were sent to the test sites. The pilot survey and the follow-ups generated a fifty-percent response rate. Both the responses and the rate indicated that additional classifications and simplifications were needed. All of these six cities were added to the main list as part of the general survey.

Second, technical questions were categorized and some items were dropped from the questionnaire to make it easier for the respondents to complete. Questions about the utilization of census data and geographic base file maintenance were added later. The final draft was reviewed by the dissertation committee and was sent with personal letters

from the project director and the Director of the Center for Urban Studies to six more cities. Comments and suggestions were requested. This pilot survey had a 100-percent response rate after a follow-up letter from the project director was received. The second group of pilot survey responses also suggested some changes in the order of the questions. The questionnaire was revised accordingly.

### Survey Implementation

Surveys and follow-ups were mailed to all directors in the sample according to the schedule in Table II. Each survey mailing included a personalized cover letter and a business reply envelope. The costs of mailing, copying, paper, and envelopes were covered by the Center for Urban Studies.

TABLE II  
MAILING SCHEDULE FOR SURVEYS

<u>Date (1991)</u>	<u>Type of Mailing</u>
February 4	Initial mailing of survey
February 18	First follow-up letter reminder to all directors
March 25	Second follow-up letter with questionnaire to non-respondents

With the exception of the suggested time frame, these mailings followed the strategy suggested by Bainbridge (1989) and Dillman (1978). They suggested a period of up to seven weeks to finish mailing, including three follow-ups with first-class mail and telephone calls. There were three ways that the survey departed from their strategy. First, bulk-rate mail was used for the main survey and the two follow-ups. Second, questionnaires were sent to directors who may not have been the most appropriate persons to complete them. Third, more than one person might have been needed to complete the questionnaire. All of the above factors required more time than Bainbridge and Dillman estimated.

### Survey Response

Table III reports the sequence of mailings and returns. It should be noted that the first follow-up generated thirty-two phone calls and letters between February 26 and March 22. Two of the callers had questions, while all others requested replacement questionnaires. These calls indicated that the follow-up letters were not being forwarded to the same people who had received the questionnaires. Replacements were mailed within two days by first-class mail. Four-fifths of the replacement questionnaires were returned, with three instances in which the questionnaire was completed twice by the same department. Responses from the callers requesting replacements were returned more quickly than the main survey, with the replacements taking only about ten days. This faster return rate may have occurred because first-class mail was used, and it was verifiable because the researcher knew the names of the callers. In all, 280 (66.5%) questionnaires were returned and a third follow-up was not necessary.

TABLE III  
COMPLETED SURVEYS RETURNED EACH WEEK

<u>Returns by Week</u>	<u>Number</u>	<u>Percentage</u>	<u>Cumulative Percentage</u>
Surveys Mailed February 4			
Week 2 (Feb. 11-17)	3	0.7%	0.7%
Follow-up Letters Mailed February 18			
Week 3 (Feb. 18-24)	17	4.0%	4.7%
Week 4 (Feb. 25-March 3)	67	15.9%	20.6%
Week 5 (March 4-10)	41	9.8%	30.4%
Week 6 (March 11-17)	28	6.7%	37.1%
Week 7 (March 18-24)	23	5.5%	42.6%
Follow-up Surveys Mailed March 25			
Week 8 (March 25-31)	7	1.6%	44.2%
Week 9 (April 1-7)	2	0.5%	44.7%
Week 10 (April 8-14)	57	13.5%	58.2%
Week 11 (April 15-21)	14	3.3%	61.5%
Week 12 (April 22-28)	5	1.2%	62.7%
Week 13 or later	16	3.8%	66.5%
No response	141	33.5%	100.0%
Total	421	100.0%	100.0%

Two hundred sixty of the returned questionnaires were coded and entered into a Mac II microcomputer using Microsoft Excel software. The remaining twenty questionnaires were incomplete or respondents indicated not using GIS. All six returned questionnaires used in the second pilot study were entered into the database -- which made the total returns 266 and the total sample 427. It should be noted that the mean areas and populations of responding and all cities were very close, indicating that the returned sample is representative of the general population (see Appendix B, Table B-I).

## SECONDARY DATA COLLECTION

For each responding department, data on city population size were obtained from the 1990 census, and data on area from the County/City Data Book (1988). Some city areas were not available in this source because the townships were not defined as municipal governments. This unavailability of data causes the number of cities to vary in the analyses.

## DATA ANALYSIS

Two types of analyses are used. First, descriptive statistics are used to discuss technical, organizational, and institutional environments. Also, cluster analysis is used to differentiate between responding cities using the above environments. Second, nonparametric analysis is used to examine the relationships between advancements in geocoding methods and data use, spatial resolution, and data volume; between spatial resolution and data volume and use; and between advancements in geocoding methods and city areas, city populations, and department budgets.

Chapter IV presents the findings, and Chapter V discusses whether the research hypotheses were supported.

## CHAPTER IV

### DESCRIPTIVE RESULTS

One purpose of this study was to identify and compare planning departments in terms of their technical, organizational, and institutional environments. Results appearing in this chapter are divided into four sections. The technical section identifies the respondents as automated or manual departments, discusses spatial resolution and types of data for automated departments, compares automated and manual departments in terms of hardware configurations and application software, and discusses forms of GBFs and types of digital map bases used for displaying in automated departments. The organizational section makes comparisons among locations of GIS, locations of GIS in the metropolitan area, and GBFs' maintenance locations. The institutional section compares automated and manual departments using the factors of types and numbers of full-time employees and types and hours of training (City area, population, and department budget are analyzed in the following chapter). Finally, in the grouping section, respondents were grouped into five clusters based on their technical, organizational, and institutional environments.

Since respondents were asked to specify all categories that applied in most questions, percentages in the tables represent frequency of responses for each category for each type of department. These percentages were calculated by multiplying each frequency by 100 and dividing it by the sample size. For example, 136 manual departments used a "look up on map" method to reference addresses to geographic areas. This number was multiplied by 100 and divided by the sample size (154), which

produced the eighty-eight percent. Twelve percent of manual departments did not use this method and were not reported in the table (see Table IV). The reasons for calculating and reporting column percentages for each category were: 1) the sample sizes for automated and manual departments were different, 2) respondents were asked to specify all categories that applied, and 3) responses for each category needed to be compared and discussed. Therefore, percentages added up to more than 100 percent, indicating that respondents specified more than one category. This means that two or more categories were specified, and usually categories with higher percentages are the most commonly combined categories.

## TECHNICAL ENVIRONMENT

This section shows how planning departments geocoded street addresses; what areal units were used; and what types of data, geocoding applications, hardware configurations, software programs, Geographic Base Files, and digital base maps were used for geocoding and displaying.

### Types of Departments

The returned questionnaires were divided into two groups, depending on the responses to the first question: "Which of the following geocoding methods are used to assign street addresses to geographic locations?" Respondents who specified only the first two categories, "look up on map" and "look up in directory" (both manual geocoding methods), were designated manual departments, while respondents who specified any automated categories were designated automated departments. One hundred fifty-four (58%) of the planning departments used only manual geocoding methods to assign data to geographic locations, while 112 (42%) used one or more of the automated geocoding methods listed in the questionnaire to assign data to geographic locations automatically.

This result does not mean that manual departments do not use automated systems and GIS. They might have these systems in other departments or use them for purposes other than geocoding. Also, automated departments use manual geocoding methods in addition to the automated methods they specified. But for clarity and discussion purposes, all departments are designated as either manual or automated.

Manual Departments. Manual geocoding methods are still a common practice among planning departments. Fifty-eight percent of those returning questionnaires indicated that they geocoded data manually. Table IV shows the percentages of manual departments using each of the two manual geocoding methods. Eighty-eight percent of manual departments use the “look up on map” method, while only forty-one percent use the “look up in directory” method. “Look up on map” (direct) geocoding is more common than “look up in directory” (indirect) geocoding among manual departments. The “look up on map” method is used primarily to map data for visual analysis.

TABLE IV  
PERCENTAGES OF MANUAL DEPARTMENTS USING  
MANUAL GEOCODING METHODS

<u>Methods of Geocoding</u>	<u>Manual Departments</u>
	N = 154
Look up on map	88%
Look up in directory	41%

N = Number of Departments

Source: Question 1

Note: Percentages sum up to more than 100% because respondents specified more than one category.

Automated Departments. Forty-two percent of the responding departments indicated that they geocoded street addresses automatically. Table V shows five automated methods and the percentages of automated departments using each method. Batch and interactive processing with tabular output and geocoding within a GIS context



were the most common methods used for address geocoding, and more than half of the automated departments geocoded addresses within a GIS context. The transfer of geocoded data to a Thematic Mapping Program and geocoding within a specialized system were used by only nineteen and eleven percent of automated departments, respectively. Only six percent used other methods of geocoding, and a few departments received geocoded data from other departments.

TABLE V  
PERCENTAGES OF AUTOMATED DEPARTMENTS USING  
AUTOMATED GEOCODING METHODS

<u>Methods of Automated Geocoding</u>	<u>Automated Departments</u>
	N = 112
Batch processing with tabular output	46%
Interactive processing with tabular output	42%
Transfer of geocoded data to a Thematic Mapping Program	19%
Geocoding within a GIS	56%
Geocoding within specialized systems	11%
Other	6%

N = Number of Departments

Source: Question 1

Note: Percentages sum up to more than 100% because respondents specified more than one category.

It seems that a combination of two or three methods of automated geocoding was used by a number of planning departments. Geocoding within a GIS, combined with batch and interactive geocoding, were the most common combinations of geocoding methods. A small number of cities transferred data to Thematic Mapping Programs and used specialized systems for geocoding. This probably happened because a large number of automated departments developed their own systems that were not compatible with the new systems available outside their departments. It is possible to use one or more

of the batch, interactive, or geocoding within a GIS context, but it is not possible to use any of these three methods with a transfer of geocoded data or a specialized system in the same computer environment.

The batch and interactive modes of geocoding programs from the 1960s and 1970s are still used for geocoding because of their ready availability. Also, geocoding within a GIS context is commonly available among U.S. cities. Generally, a small number of U.S. cities geocode street addresses automatically, and a smaller number (approximately 10%) geocode addresses within a GIS context.

#### Data Volume and Use

Automated geocoding and GIS technology are used to conduct advanced geocoding applications using urban data files and maps. This section considers the following questions: 1) What sizes of geographic areas are used for geocoding? 2) What types of urban data are used for geocoding? 3) What types of applications are being used? Data about resolution and types of data are available only for automated departments, because manual departments did not answer these questions (see questions 4 and 5). The data on spatial resolution and types of data follow. Quantity of spatial resolution, quantity of data transactions, and frequency of updates are computed to determine data volume.

Spatial Resolution. Automated departments use many types of areal units for geocoding. The author grouped them into six categories and arranged them by size, as presented in Table VI. Some respondents specified up to five additional areal units. None of the additional units were smaller than ownership parcels.

Three-fourths of the automated departments used planning areas, census tracts, and ownership parcels as areal units for geocoding. About one-half of the automated departments used traffic zones, city block groups, and city blocks. Planning areas, census

tracts, and ownership parcels were the most common types of spatial resolution used for geocoding, indicating that automated departments used the largest and smallest geographic units for geocoding. Large geographic areas may be used for planning purposes such as urban and economic development and are linked with attribute data using nominal geocodes. Ownership parcels may be used for land use planning, tax assessment, or data retrieval and are linked with attribute data using positional geocodes.

TABLE VI  
PERCENTAGES OF SPATIAL RESOLUTION BY AUTOMATED DEPARTMENTS

<u>Spatial Resolution</u>	<u>Automated Departments</u>
	N = 112
Planning areas	75%
Census tracts	81%
Traffic zones	54%
City block groups	46%
City blocks	52%
Ownership parcels	74%

N = Number of Departments

Source: Question 4

Note: Percentages sum up to more than 100% because respondents specified more than one category.

Data Types. Table VII shows that eight different types of data were used.

Building permits/inspections and land parcel data are the most frequently geocoded data in planning departments. Housing inspections and employer/employee locations are less frequently used. Less than ten percent of automated departments use vital statistics data, retail sales data, auto registration data, and travel data. This table shows that planning departments use housing and building permits, employees, and land parcel data related to planning areas, census tracts, and ownership parcels as shown in the previous table.

Vital, retail, auto registration, and travel data are probably geocoded and used by other departments.

TABLE VII  
PERCENTAGES OF DATA TYPES BY AUTOMATED DEPARTMENTS

<u>Types of Data</u>	<u>Automated Departments</u>
	N = 112
Vital statistics	9%
Housing inspections	19%
Building permits/inspections	54%
Retail sales data	7%
Employer/employee locations	23%
Auto registration data	4%
Travel data	4%
Land parcel data	81%

N = Number of Departments

Source: Question 5

Note: Percentages sum up to more than 100% because respondents specified more than one category.

Questions 2 and 3 provided information about types of geocoding applications and types of census applications. In the table, the types of applications were ranked from the least advanced to the most advanced.

Geocoding Applications. Table VIII compares geocoding applications between automated and manual departments. Automated departments employed all types of applications twice as frequently as did manual departments. Although the percentages

TABLE VIII  
TYPES OF GEOCODING APPLICATIONS BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Geocoding Applications</u>	<u>Automated Departments</u>	<u>Manual Departments</u>
	N = 112	N = 154
Tabulating addresses by statistical area without mapping the results	70%	37%
Aggregating addresses to statistical areas and mapping	68%	36%
Matching addresses to coordinates and/or network nodes	50%	12%
Linking coordinate-based addresses to modeling and other applications programs	29%	5%

N = Number of Departments

Source: Question 2

Note: Percentages sum up to more than 100% because respondents specified more than one category.

declined from less-advanced utilization to the most-advanced utilization for both types of users, the percentages for the most-advanced categories of utilization were four to five times higher for automated departments than for manual departments. Higher percentages of automated departments are able to apply advanced geocoding applications such as mapping incidents and input data to other analytical or modeling programs.

Census Applications. Table IX compares census applications between automated and manual departments. The utilizations were ordered in six categories from the least advanced to the most advanced. This table shows that the percentages of manual departments that extract and tabulate data from printed reports and map tract data manually are higher than the percentages of automated departments. On the other hand, the percentages of automated departments that map tract data, tabulate block and block group data from summary tapes, map block and block group data from summary tapes, and develop usage rates with geocoded data as the nominator and census data as the denominator (higher-level applications) are higher than the corresponding percentages of manual departments. The difference can be explained as a result of the availability of

TABLE IX  
TYPES OF CENSUS APPLICATIONS BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Census Applications</u>	<u>Automated Departments</u> N = 112	<u>Manual Departments</u> N = 154
Extraction of tabular data from printed reports	79%	84%
Manual mapping of tract data	49%	64%
Computer mapping of tract data	54%	15%
Tabulation of block and block group data from summary tapes	55%	22%
Mapping of block and block group data from summary tapes	41%	8%
Development of usage rates with geocoded data as the numerator and census data as the denominator	15%	1%

N = Number of Departments

Source: Question 3

Note: Percentages sum up to more than 100% because respondents specified more than one category.

geocoding and GIS technology in automated departments. Mapping and tabulating census data automatically at the block and ownership parcel levels and developing complex applications are the major advantages of using advanced geocoding and GIS technology in automated planning departments.

### Hardware Configurations

Table X compares types of hardware configurations between automated and manual departments. As expected, automated departments have more hardware of all types than manual departments. Terminals hooked to a mainframe in a DP department and microcomputers are the most common types of hardware available in both automated and manual departments. While nearly one-fifth of automated departments have terminals hooked to a minicomputer in another department or in a planning department, fewer than one-tenth of manual departments do. Networked workstations are the third most common form of configuration in both automated and manual departments. Both types of departments are investing heavily in this new hardware technology.

TABLE X  
TYPES OF HARDWARE CONFIGURATIONS BY AUTOMATED AND  
MANUAL DEPARTMENTS

<u>Hardware Configurations</u>	<u>Automated Departments</u> N = 112	<u>Manual Departments</u> N = 154
Terminals to a mainframe in a DP department	69%	57%
Terminals to a minicomputer in another department	20%	10%
Terminals to a minicomputer in planning department	13%	5%
Networked workstations	40%	31%
Microcomputers	79%	67%

N = Number of Departments

Source: Question 9

Note: Percentages sum up to more than 100% because respondents specified more than one category.

To examine the date the hardware was introduced, Table XI compares the average number of years since the hardware was introduced for each type of hardware configuration between automated and manual departments. As expected, automated departments introduced all types of hardware earlier than did manual departments, except for terminals hooked to a minicomputer in another department. The average number of years since the introduction of hardware in automated departments was six years, compared to five years for manual departments. It seems that manual departments relied on other departments for computing services and adopted the technology only recently. However, both automated and manual departments introduced networked workstations in approximately the last three years. These data illustrate a shift in the use of hardware, from mainframes to microcomputers to networked workstations.

TABLE XI

AVERAGE NUMBER OF YEARS SINCE EACH TYPE OF HARDWARE CONFIGURATION WAS INTRODUCED BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Hardware Configurations</u>	<u>Average Number of Years</u>			
	<u>Automated Departments</u>		<u>Manual Departments</u>	
	<u>Mean</u>	<u>N</u>	<u>Mean</u>	<u>N</u>
Terminals to a mainframe in a DP department	9.91	70	7.50	80
Terminals to a minicomputer in another department	5.09	22	5.85	14
Terminals to a minicomputer in planning department	5.76	13	2.87	8
Networked workstations	3.70	44	3.53	45
Microcomputers	6.44	83	5.45	100

N = Number of Departments

Source: Question 9

Table XII compares the average number of seats for each type of hardware configuration. As expected, the average numbers of terminals hooked to a mainframe in a DP department, terminals hooked to a minicomputer in another department, terminals hooked to a minicomputer in a planning department, and microcomputers were higher for

automated departments than for manual departments. The average number of seats for automated departments was double or triple the number for manual departments. Manual departments have a higher average number of networked workstations than do automated departments. It appears that manual departments are adopting new hardware technology at a later time, but with current technology.

TABLE XII  
AVERAGE NUMBER OF SEATS FOR EACH TYPE OF HARDWARE CONFIGURATION  
BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Hardware Configurations</u>	<u>Average Number of Seats</u>			
	<u>Automated Departments</u>		<u>Manual Departments</u>	
	<u>Mean</u>	<u>N</u>	<u>Mean</u>	<u>N</u>
Terminals to a mainframe in a DP department	12.59	72	9.96	76
Terminals to a minicomputer in another department	15.12	19	3.56	16
Terminals to a minicomputer in planning department	30.53	13	5.12	8
Networked workstations	9.05	39	10.76	46
Microcomputers	13.32	83	7.48	96

N = Number of Departments  
Source: Question 9

### Application Software

Table XIII compares types of application software available in automated and manual departments. Office automation systems were dominant in all departments. By definition, all types of geocoding systems were dominant in automated departments. In addition, automated departments used the computer for geocoding applications at a higher rate than manual departments. The automated departments were moving heavily into GIS, while manual departments were moving more rapidly into CADD. Manual departments may have been interested in map automation software rather than geographic



and spatial analysis. Another possibility is that manual departments did not realize the advantages of GIS and chose CADD due to its distinctive map production, or due to the departments' lack of resources.

TABLE XIII

## TYPES OF APPLICATION SOFTWARE BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Software</u>	<u>Automated Departments</u>	<u>Manual Departments</u>
	N = 112	N = 154
Office automation	95%	95%
Geocoding systems	52%	7%
Thematic Mapping	35%	13%
Computer-Aided Design and Drafting (CADD)	47%	27%
Geographic Information Systems (GIS)	63%	16%

N = Number of Departments

Source: Question 10

Note: Percentages sum up to more than 100% because respondents specified more than one category.

Table XIV compares the average number of years since the software was introduced for each type of application software. Automated departments introduced all types of software earlier than did manual departments. The average number of years

TABLE XIV

## AVERAGE NUMBER OF YEARS SINCE EACH TYPE OF APPLICATION SOFTWARE WAS INTRODUCED BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Software</u>	<u>Average Number of Years</u>			
	<u>Automated Departments</u>		<u>Manual Departments</u>	
	<u>Mean</u>	<u>N</u>	<u>Mean</u>	<u>N</u>
Office automation	7.76	101	6.41	131
Geocoding systems	6.96	56	4.80	10
Thematic Mapping	5.93	38	5.10	19
Computer-Aided Design and Drafting (CADD)	5.66	50	3.22	40
Geographic Information Systems (GIS)	3.94	68	3.39	23

N = Number of Departments

Source: Question 10

ranged from five to seven years for automated departments, compared to three to six years for manual departments, except for GIS software. The average number of years since GIS had been introduced was about four years for automated departments and three and one-half years for manual departments. Manual departments introduced GIS software in a very small number of departments in the last three years or so as a result of its availability, affordable cost, and ease of operation.

#### Geographic Base File (GBF)

Automated departments were asked to specify the types of digital forms of Geographic Base Files (GBFs) used in their departments. Table XV shows the use of these types. More than half of the automated departments used directories of street names and segment-based files such as DIME and TIGER files. Cross-street indices were used by only nineteen percent and address libraries by thirty-nine percent of automated departments. Fifteen percent used parcel files, assessor tax data, AutoCAD street files, zoning and subdivision databases, traffic zones, and street coverages. Segment-based files and street directories were the most common types of GBFs used for geocoding due to their availability at low costs, easy operation, and user satisfaction.

TABLE XV  
TYPES OF GEOGRAPHIC BASE FILES BY AUTOMATED DEPARTMENTS

<u>Forms of Geographic Base Files</u>	<u>Automated Departments</u>
	N = 112
Directory of street names and address ranges and associated areal unit codes	51%
Cross-street index	19%
Segment-based files (e.g., DIME/TIGER)	58%
Address library (universe of all addresses in your jurisdiction)	39%
Others	15%

N = Number of Departments

Source: Question 6

Note: Percentages sum up to more than 100% because respondents specified more than one category.

### Digital Map Base Used for Display

Automated departments were also asked to specify types of digital map bases used for displaying data. Table XVI provides the results. Twenty-one percent did not map address data, and only thirty-seven percent mapped data by tract/neighborhood boundaries. Very few of the automated departments used a scanned map image background with TIGER. Twenty-two percent used other digital maps such as scanned roads and hydrography, boundary files, and ETAK maps. About one-half of automated departments used a DIME/TIGER file or vector parcel layer with street right-of-way, since these files are the most useful and relatively inexpensive map bases.

TABLE XVI  
TYPES OF DIGITAL MAP BASE USED FOR DISPLAYING DATA BY  
AUTOMATED DEPARTMENTS

<u>Digital Map Base Use</u>	<u>Automated Departments</u>
	N = 112
Not mapped	21%
Tract/neighborhood boundaries	37%
DIME/TIGER/street-centerline file	49%
Scanned map image background with TIGER	5%
Vector parcel layer with street right-of-way	46%
Others	22%

N = Number of Departments

Source: Question 8

Note: Percentages sum up to more than 100% because respondents specified more than one category.

### ORGANIZATIONAL ENVIRONMENT

In this section, organizational issues are discussed. Questions about GIS locations, what other organizations in the metropolitan area use these systems, and who maintains them will be answered in the following sections.

### City Organization

The key factor for GIS location was whether the systems were centralized or decentralized. A centralized system (one central unit) was one that was located in one department in the city and served all other related departments. A decentralized system was one that was located in more than one department, with each unit serving the needs of its department only and not supporting other departments.

Table XVII shows that about two-thirds of the automated departments had centralized units, compared to about one-fifth of the manual departments. This may be because automated departments were capable of operating and maintaining their systems. As a result, automated planning departments can share more geographic data with other departments in the city. The domination of centralized units in automated departments can be explained by the fact that automated departments adopted computer systems such as the mainframe a long time ago, whereas manual departments have adopted decentralized units very recently as a result of the availability of new computer systems such as workstations and microcomputers.

TABLE XVII  
CENTRALIZED AND DECENTRALIZED GIS UNITS BY AUTOMATED AND  
MANUAL DEPARTMENTS

<u>Types of Units</u>	<u>Automated Departments</u>	<u>Manual Departments</u>
	N = 112	N = 154
Centralized units	62%	22%
Decentralized units	38%	78%
Total	100%	100%

N = Number of Departments  
Source: Question 13

Table XVIII compares the use of centralized and decentralized units by city departments between automated and manual departments. The table shows that the percentages of city departments using centralized and decentralized units indicated by

automated planning departments were higher than the corresponding percentages of manual planning departments. For example, thirty-seven percent of automated departments indicated that their public works departments used centralized units, compared to only thirteen percent of manual departments.

TABLE XVIII  
CITY DEPARTMENTS WITH CENTRALIZED AND DECENTRALIZED GIS UNITS BY  
AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Units</u>	<u>Automated Departments</u> N = 112	<u>Manual Departments</u> N = 154
Departments with centralized units		
Public works	37%	13%
Data processing	30%	10%
Public safety	27%	4%
Planning	45%	13%
Transportation/traffic	28%	4%
Other departments	35%	11%
Departments with decentralized units		
Public works	16%	12%
Data processing	8%	5%
Public safety	15%	7%
Planning	31%	10%
Transportation/traffic	10%	5%
Other departments	18%	7%

N = Number of Departments

Source: Question 13

Note: Percentages sum up to more than 100% because respondents specified more than one category.

The percentages of city departments using centralized units (indicated by planning departments) were higher than the percentages of city departments with decentralized units. For example, forty-five percent of automated departments indicated that their planning departments used centralized units, compared to thirty-one percent indicating decentralized units. Centralized units were common among cities with automated planning departments, compared to cities with manual departments. Automated departments could share data, and all city departments could select and use the resolution

and volume required to satisfy their requirements. For example, data geocoded by ownership parcels or point locations in public works or transportation planning departments could be retrieved and aggregated by larger areas for the purpose of mapping and planning. In a decentralized environment, geographic data sharing was difficult unless separate units were networked and were compatible.

#### Locations of GIS in the Metropolitan Area

Table XIX compares locations of GIS in the metropolitan area. This question included three different types of organizations in the metropolitan area, as shown in the table. Although the category “none” was included in the question, many respondents did not answer the question and their omissions were considered missing responses. Perhaps the respondents who did not answer this question did not know whether any other organizations in the area had GIS, so they were reluctant to specify “none.”

TABLE XIX

#### LOCATIONS OF GIS IN OTHER ORGANIZATIONS IN THE METROPOLITAN AREA BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Organizations</u>	<u>Automated Departments</u>	<u>Manual Departments</u>
	N = 112	N = 154
None	12%	32%
City	58%	19%
County	48%	42%
Metropolitan planning agency	36%	22%
Other organizations	20%	20%

N = Number of Departments

Source: Question 14

Note: Percentages sum up to more than 100% because respondents specified more than one category.

The percentages of automated departments using GIS in cities, counties, and metropolitan planning agencies were higher than the corresponding percentages of manual departments. It appears that when GIS was available to automated departments in

a city, GIS was more readily available in the other organizations in the metropolitan area. Note that the percentage of manual departments specifying the category “none” was higher than the corresponding percentage of automated departments. This difference indicates that, with manual departments, many of the surrounding organizations did not have GIS.

### Geographic Base File Maintenance

Table XX provides information about different methods of maintaining GBFs. Sixty-six percent of automated departments maintained their own files. Twenty percent of automated departments shared maintenance within cities and counties, only seven percent shared maintenance within functions, and five percent shared maintenance metropolitan-wide for all functions. Experience of automated departments with hardware and software programs helped them to be more independent in maintaining and operating their systems.

TABLE XX  
MAINTENANCE LOCATIONS OF GEOGRAPHIC BASE FILES BY  
AUTOMATED DEPARTMENTS

<u>GBF Maintenance Locations</u>	<u>Automated Departments</u>
	N = 112
Each agency/department maintains its own files	66%
Shared maintenance within cities/counties	20%
Shared maintenance metropolitan-wide within functions (e.g., public safety, schools, planning, etc.)	8%
Shared maintenance metropolitan-wide for all functions	6%
Total	100%

N = Number of Departments

Source: Question 7

Note: Percentages sum up to more than 100% because respondents specified more than one category.

## INSTITUTIONAL ENVIRONMENT

This section discusses and compares automated and manual departments in terms of types and numbers of full-time employees and types and hours of training. Other institutional variables such as city area, city population, and department budget are analyzed in the following chapter.

### Full-Time Employees

Table XXI compares types of full-time employees available in automated and manual departments. Automated departments had higher percentages of all types of employees than did manual departments. Users/analysts and operators/technicians were the most common types of employees in both types of departments. Systems support and database administrators were less common. It appears that both types of departments are using employees primarily to operate and analyze data. The use of employees to support and maintain systems and data will likely increase as investments in systems and data increase.

TABLE XXI  
TYPES OF FULL-TIME EMPLOYEES BY AUTOMATED AND  
MANUAL DEPARTMENTS

<u>Types of Employees</u>	<u>Automated Departments</u>	<u>Manual Departments</u>
	N = 112	N = 154
Users/analysts	65%	23%
Systems support	31%	7%
Database administrators	29%	12%
Operators/technicians	45%	20%

N = Number of Departments

Source: Question 11

Note: Percentages sum up to more than 100% because respondents specified more than one category.



Table XXII compares the average numbers of different types of employees between automated and manual departments. It was expected that automated departments would have larger numbers of all types of employees. This expectation was validated only for operators/technicians. The large number of employees in manual departments can be explained as a result of these departments reporting all employees dealing with computer services in all departments in the city, especially since most of them used decentralized units. Automated departments may have reported the exact numbers of all types of full-time employees dealing with GIS and geocoding systems in their departments, especially since most of them used centralized units.

TABLE XXII

AVERAGE NUMBER OF EACH TYPE OF FULL-TIME EMPLOYEES BY  
AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Employees</u>	<u>Average Numbers of Full-Time Employees</u>			
	<u>Automated Departments</u>		<u>Manual Departments</u>	
	<u>Mean</u>	<u>N</u>	<u>Mean</u>	<u>N</u>
Users/analysts	4.25	72	5.31	32
Systems support	1.45	30	1.68	11
Database administrators	1.25	33	1.40	17
Operators/technicians	2.21	51	1.91	30

N = Number of Departments

Source: Question 11

### Types of Training

Table XXIII compares types of training between automated and manual departments. Except for community college/university courses, automated departments used all available types of training more often than did manual departments. Ten percent of manual departments trained their employees in community college/university courses, compared to nine percent of automated departments. In-house and vendor training were the types of training most frequently used by both automated and manual departments. These two types of training are probably the most effective and available to users.

TABLE XXIII  
TYPES OF TRAINING BY AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Training</u>	<u>Automated Departments</u>	<u>Manual Departments</u>
	N = 112	N = 154
In-house training	60%	20%
Consultant training	15%	9%
Short courses	16%	4%
Community college/university courses	9%	10%
Vendor training	33%	13%

N = Number of Departments

Source: Question 12

Note: Percentages sum up to more than 100% because respondents specified more than one category.

Table XXIV compares the average number of hours of different types of training between automated and manual departments. As expected, the research found that automated departments provided their employees with more training than did manual departments. The average numbers of hours for all types of training were higher for automated departments than for manual departments, except for short course training. Automated departments trained their employees an average of one hundred training hours per year by in-house, consultant, and vendor training, compared to an average of fifty

TABLE XXIV  
AVERAGE NUMBER OF HOURS FOR EACH TYPE OF TRAINING BY  
AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Training</u>	<u>Average Numbers of Hours of Training</u>			
	<u>Automated Departments</u>		<u>Manual Departments</u>	
	<u>Mean</u>	<u>N</u>	<u>Mean</u>	<u>N</u>
In-house training	83.11	45	49.47	21
Consultant training	93.78	14	22.92	13
Short courses	34.92	14	37.60	5
Community college/university courses	54.11	9	22.66	5
Vendor training	105.55	27	44.00	13

N = Number of Departments

Source: Question 12

hours per year for manual departments. GIS required more training time than any other geocoding system. Also, as expected, the number of hours of training by short courses was the smallest average for both departments because these courses were offered mostly for one to five days. Manual departments had a slightly higher average number of hours of training by short courses, suggesting that CADD systems required short periods of training time. GIS required more specialized training for longer periods of time for both in-house and vendor training.

The following section will use the above environments to segregate responding cities into five groups. Each group shares similar characteristics which provide detailed information about each group.

#### PLANNING DEPARTMENT GROUPING

Previous sections identified and discussed planning departments in terms of their technical, organizational, and institutional environments, but the diversity of planning departments leaves one wondering if there might be identifiable groups with different technical, organizational, and institutional environments. K-means cluster analysis was used to answer this question. All responding cities were included in the analysis. Two to nine clusters were identified using the variables listed in Table XXV.

The process used to identify these clusters, K-means cluster analysis, uses an iterative procedure to form a specified number of groups, with large differences between groups and small differences within groups (Wilkinson, 1989). The grouping depends on the variables used, the way the variables are scaled, and the number of clusters specified. Five clusters are used, given the previous understanding of the data.

TABLE XXV  
DEFINITION OF VARIABLES USED IN K-MEANS CLUSTER ANALYSIS

<u>Variable</u>	<u>Definition</u>
AREA	City area in square miles
DECENTRALIZED	Number of departments using decentralized units
INCOMEX	In-house computing experience in years using workstations
GEOMETHODS	Methods of geocoding used to assign street addresses to geographic locations
POPULATION	City population
OFFICE	Office automation software experience in years (e.g., wordprocessing and spreadsheets)
CENTRALIZED	Number of departments using centralized units
EXCOMEX	External computing experience in years using mainframe and minicomputers
MICRO	Microcomputer experience in years
DATAUSE	Types of geocoding applications
CADD	Computer-aided Design and Drafting software experience in years (e.g., AutoCAD)
GIS	GIS software experience in years (e.g., ARC/INFO)
THEMATIC	Thematic Mapping software experience in years (e.g., ATLAS)
GEOCODING	Geocoding software experience in years (e.g., ADDMATCH and UNIMATCH)
BUDGET	Planning department annual budget

Table XXVI presents summary statistics for the cluster analysis as well as for each cluster. The summary statistics show that area and decentralized units were the most important variables separating groups, followed by internal hardware experience and methods of geocoding. All other variables were also statistically significant at the ninety-five percent level of confidence, except department budget. All variables were standardized so that larger-range variables would not overwhelm smaller-range variables. Real values are included in Appendix B, Table B-II.

The statistics for each cluster reveal the differences which helped determine the clusters. Area, decentralized systems, minicomputer and workstation experience in years, and methods of geocoding were the important variables for distinguishing groups. The cluster of non-computer-oriented departments had the lowest averages of all variables. They use manual geocoding; conduct elementary geocoding applications; have small populations, areas, and budget sizes; have less experience with hardware and software programs; and do not yet have known organizational structures.

TABLE XXVI  
CLUSTER ANALYSIS RESULTS FOR PLANNING DEPARTMENTS

<u>Summary Statistics (N=266)</u>				<u>Degrees of</u>
<u>Variables</u>	<u>Between SS</u>	<u>Within SS</u>	<u>F-ratio</u>	<u>Freedom</u>
AREA	182.104	72.897	156.757	4, 251
DECENTRALIZED	135.475	110.988	79.036	4, 259
INCOMEX	100.700	131.478	49.401	4, 258
GEOMETHODS	96.234	151.406	39.089	4, 246
POPULATION	25.483	62.367	26.354	4, 258
OFFICE	72.603	157.287	26.080	4, 266
CENTRALIZED	66.802	187.306	23.093	4, 259
EXCOMEX	51.000	186.667	17.622	4, 258
MICRO	49.769	130.807	16.741	4, 176
DATAUSE	43.564	144.007	14.067	4, 186
CADD	33.601	54.546	12.936	4, 84
GIS	11.912	35.389	6.985	4, 83
THEMATIC	14.212	29.650	6.111	4, 51
GEOCODING	15.828	44.583	5.236	4, 59
BUDGET	0.996	26.752	1.955	4, 210
<u>Cluster of Non-Computer-Oriented</u>				
<u>Departments (N=152)</u>				
<u>Variables</u>	<u>Minimum</u>	<u>Mean</u>	<u>Maximum</u>	<u>Standard Dev.</u>
AREA	-0.63	-0.25	1.73	0.38
DECENTRALIZED	-0.56	-0.36	1.05	0.43
INCOMEX	-0.55	-0.25	1.89	0.53
GEOMETHODS	-0.83	-0.55	1.47	0.58
POPULATION	-0.48	-0.25	0.76	0.25
OFFICE	-2.06	-0.44	1.71	0.87
CENTRALIZED	-0.61	-0.30	2.58	0.68
EXCOMEX	-0.55	-0.25	1.89	0.53
MICRO	-1.77	-0.41	2.20	0.82
DATAUSE	-1.38	-0.48	1.48	0.81
CADD	-0.96	-0.49	0.11	0.30
GIS	-0.78	-0.33	0.61	0.38
THEMATIC	-1.31	-0.45	0.09	0.39
GEOCODING	-1.01	-0.47	0.42	0.40
BUDGET	-0.37	-0.12	2.45	0.36

TABLE XXVI

DEFINITION OF VARIABLES USED IN K-MEANS CLUSTER ANALYSIS  
(Continued)

<u>Cluster of CADD Mapping Departments (N=5)</u>				
<u>Variables</u>	<u>Minimum</u>	<u>Mean</u>	<u>Maximum</u>	<u>Standard Dev.</u>
AREA	-0.51	0.08	1.19	0.60
DECENTRALIZED	-0.56	-0.24	0.24	0.39
INCOMEX	2.93	4.04	5.02	0.81
GEOMETHODS	-0.83	0.55	1.47	0.86
POPULATION	-0.45	0.61	2.36	1.02
OFFICE	0.34	1.92	3.08	0.91
CENTRALIZED	-0.61	0.45	2.04	0.95
EXCOMEX	-0.82	0.77	1.58	0.90
MICRO	-1.05	-0.05	0.76	0.65
DATAUSE	-1.38	-0.23	1.48	1.11
CADD	1.99	2.93	3.87	0.94
GIS	-0.22	1.54	2.56	1.25
THEMATIC	-0.19	0.93	2.05	0.92
GEOCODING	-0.83	-0.56	0.06	0.37
BUDGET	-0.32	-0.23	-0.13	0.10
<u>Cluster of Departments Using Geocoding on Mainframe (N=58)</u>				
<u>Variable</u>	<u>Minimum</u>	<u>Mean</u>	<u>Maximum</u>	<u>Standard Dev.</u>
AREA	-0.61	-0.10	2.18	0.48
DECENTRALIZED	-0.56	-0.35	1.05	0.44
INCOMEX	-0.55	-0.17	2.58	0.71
GEOMETHODS	-0.83	0.77	1.47	0.99
POPULATION	-0.47	0.02	1.76	0.53
OFFICE	-1.03	0.72	2.40	0.74
CENTRALIZED	-0.61	0.89	2.58	1.21
EXCOMEX	-0.82	0.74	4.59	1.31
MICRO	-1.05	0.85	4.01	0.88
DATAUSE	-1.38	0.43	1.48	1.00
CADD	-0.69	0.50	4.94	1.23
GIS	-0.78	0.09	2.56	0.81
THEMATIC	-1.03	0.52	3.17	0.95
GEOCODING	-0.83	0.61	2.57	1.09
BUDGET	-0.34	-0.07	1.14	0.30

TABLE XXVI

DEFINITION OF VARIABLES USED IN K-MEANS CLUSTER ANALYSIS  
(Continued)

<u>Cluster of Large-Budget Departments in Large-Area Cities (N=10)</u>				
<u>Variables</u>	<u>Minimum</u>	<u>Mean</u>	<u>Maximum</u>	<u>Standard Dev.</u>
AREA	1.97	4.15	7.76	1.60
DECENTRALIZED	-0.56	0.69	4.26	1.47
INCOMEX	-0.55	0.19	2.93	1.19
GEOMETHODS	-0.83	0.23	1.47	1.08
POPULATION	-0.46	1.30	3.13	1.10
OFFICE	-1.03	-0.08	1.37	0.75
CENTRALIZED	-0.61	0.22	2.58	1.23
EXCOMEX	-0.67	0.20	1.13	0.53
MICRO	-1.77	-0.29	0.76	0.75
DATAUSE	-1.38	0.05	1.48	0.82
CADD	-0.96	-0.36	0.11	0.40
GIS	-0.78	-0.13	0.89	0.66
THEMATIC	-1.31	-0.75	-0.19	0.43
GEOCODING	-0.01	0.30	2.92	1.85
BUDGET	-0.29	0.12	0.85	0.35
<u>Cluster of Departments New to GIS and Geocoding (N=41)</u>				
<u>Variables</u>	<u>Minimum</u>	<u>Mean</u>	<u>Maximum</u>	<u>Standard Dev.</u>
AREA	-0.65	0.00	1.53	0.54
DECENTRALIZED	0.24	1.60	4.26	1.13
INCOMEX	-0.55	0.44	3.28	1.04
GEOMETHODS	-0.83	0.68	1.47	0.95
POPULATION	-0.47	0.18	2.18	0.73
OFFICE	-1.37	0.26	2.05	0.81
CENTRALIZED	-0.61	-0.38	2.58	0.59
EXCOMEX	-0.82	-0.34	1.58	0.62
MICRO	-1.37	0.26	2.05	0.81
DATAUSE	-1.38	0.61	1.48	0.77
CADD	-0.69	-0.17	0.92	0.50
GIS	-0.78	-0.35	0.33	0.29
THEMATIC	-1.03	-0.29	1.49	0.70
GEOCODING	-0.83	-0.38	1.32	0.52
BUDGET	-0.35	0.04	1.34	0.41

N = Number of Departments

The cluster of CADD mapping departments have had office automation programs for a long period of time, and have extensive experience on minicomputers and workstations; but they use elementary geocoding applications and have small budgets, moderate populations, and moderate area sizes. They have used GIS and Thematic Mapping software for a long period of time, but have used geocoding programs for a short period of time. Although these departments use CADD systems on microcomputers and workstations, the statistics show that they use centralized units and moderate to advanced geocoding methods. These departments have been introduced to the computing technology, whereas the previous cluster of departments have not yet started using it.

The cluster of departments using geocoding on mainframe computers have moderate city areas, department budgets, and populations. They have had mainframes for a short period of time, and have had workstations and microcomputers for a long period of time. They have been using office automation, geocoding, Thematic Mapping, CADD, and GIS systems for a long period of time. They use centralized units. They use advanced geocoding methods and conduct advanced planning applications.

The cluster of large-budget departments in large-area cities conduct moderate geocoding applications, and have had workstations for a long period of time and mainframes for a short period of time. They have been using CADD and GIS for a short period of time, and have been using office automation software, geocoding, and Thematic Mapping programs for a moderate period of time. They use both centralized and decentralized units, and conduct advanced geocoding applications using moderate to advanced geocoding systems.

The cluster of departments new to GIS and geocoding conduct advanced geocoding applications. They have moderate populations and budget sizes and small area sizes. They have been using workstations and microcomputers for a moderate period of



time, and have been using all types of geocoding systems for a short period of time. They have been using office automation systems for a moderate period of time, and are highly decentralized. This group is highly advanced compared to the previous ones, although they have introduced GIS hardware and software technology only in the last few years. They use a broader range of planning applications that include mapping and spatial analysis than departments with only CADD systems use.

Non-computer-oriented departments (152) are the stereotypical planning departments with low means for all variables. The CADD mapping departments (5) have had CADD systems for a long period of time, and have been using minicomputers and workstation computers for a long time. Departments using geocoding on mainframes (58) have small-area cities and have been using all types of software programs, especially office automation software, for a long time. Large-budget departments in large-area cities (10) have the largest area and population sizes, but they conduct moderate numbers of geocoding applications. They have been using workstations and microcomputers for a short period of time, and have been using office automation software programs longer than other software programs. Departments new to GIS and geocoding (41) have been using advanced geocoding applications. They have been using microcomputers and workstations for a moderate period of time. They have been using geocoding, Thematic Mapping programs, CADD, and GIS software for the last two years or so. They are highly decentralized, have moderate budget and population sizes, and have small area sizes. They use a variety of planning applications compared to other departments.

There was no clear pattern of group segregation. There was a large number of cities that were small in area, population, and department budget. These cities have started to conduct elementary geocoding applications, using manual geocoding methods. A moderate number of departments in cities that are moderate in area, population, and

department budget have been using old hardware and software for a long time. They have been using batch and interactive geocoding methods on mainframe computers. A similar number of departments in small-area cities use advanced geocoding systems and conduct a variety of planning applications.

Conversely, there was a small number of cities that were large in area, population, and department budget. They have been using old hardware and software programs. Also, a small number used CADD systems. In short, a small number of large-area cities have been using old geocoding systems for a long time, compared to a small number of small-area cities that have been using advanced geocoding systems and GIS for a short time.

## SUMMARY

Technical, organizational, and institutional environments were discussed in this chapter. The technical section identified departments that geocode data manually as manual departments and those that geocode data automatically as automated departments. Batch and interactive processing, and geocoding within a GIS context, were the most common types of geocoding specified by automated departments.

Automated departments specified building permits and land parcel data as the data used and indicated that planning areas, census tracts, and ownership parcels were the most common areal units used for geocoding. In terms of data usage, automated departments with large numbers of employees conducted advanced geocoding and census applications more than did manual departments.

Automated departments used more types of hardware and software and introduced them earlier than did manual departments. Automated departments had larger numbers of hardware systems than did manual departments, except for networked workstations, of

which manual departments had larger numbers than did automated departments. Also, more than half of the automated departments used street directories and segment-based files (e.g., DIME/TIGER) as reference bases and used segment- and parcel-based files as digital maps for displaying.

The organizational section showed that about two-thirds of automated departments used centralized units, compared to one-third of manual departments. Also, automated departments had higher percentages using decentralized units in other city departments than did manual departments. In addition, automated departments were surrounded by agencies using geocoding systems more than were manual departments. Two-thirds of automated departments maintained their GBFs in their departments, whereas the rest maintained and shared them with other departments.

The institutional section showed that automated and manual departments were compared in terms of types and numbers of full-time employees, and types and hours of training. The results showed that large percentages of automated departments used all types of employees but with smaller numbers compared to manual departments. For types of training, larger percentages of automated departments trained their employees and used all types of training than did manual departments.

Finally, cluster analysis showed that area, decentralized city organization, minicomputers and workstations, and methods of geocoding were the most important variables in separating the sample into five clusters. All other variables were also significant, except department budget. The non-computer-oriented group used manual geocoding methods and had small areas, small populations, small department budgets, and little experience with hardware systems and software programs. Another group were classified as CADD mapping departments and had small department budgets, but were experienced in minicomputer and workstation hardware and had used CADD software

programs for a long time. Departments using geocoding on mainframes have used geocoding systems for a long period of time, and were moderate in all variables. Large-budget departments in large-area cities used advanced methods of geocoding and had large population and area sizes. They had some experience in geocoding programs, but not in Thematic Mapping, CADD, and GIS. They depended on mainframes and minicomputers. Departments new to GIS and geocoding used advanced geocoding methods and started using geocoding technology only very recently. They had moderate population and budget sizes, but small area sizes. These departments conducted a broader range of planning applications than did departments using geocoding on mainframes. Some large-area cities were still using mainframe computers for geocoding, while some small cities were using advanced geocoding systems.

## CONCLUSIONS

These descriptive results helped us understand technical, organizational, and institutional environments. A small number of large cities and a large number of small cities use advanced geocoding systems. Although the majority of the cities use manual geocoding methods, a moderate number of cities have begun using automated geocoding and GIS technology. Some of these cities are small in size and have adopted the technology only in the last two or three years. As expected, departments using new technology conduct advanced planning applications and are better off than departments using manual geocoding methods. Automated geocoding and GIS technology help city planners to process, map, and analyze data using available and current information.

Departments with GIS use both large and small geographic areas, and use only two or three types of data. Seventy-five percent of automated departments that use GIS use large areas for planning and small areas for management and engineering purposes.

Fifty percent use the city block level for planning purposes. GIS users may change their analysis from the planning area and census tract levels to the city block level for mapping and planning. This study does not show how non-computer-oriented departments and departments with mainframes and minicomputers use different sizes of geographic areas, and further investigation is needed to explore this issue.

Another major finding was that departments with mainframes and minicomputers were the first to implement and use hardware and software during the 1960s and 1970s. However, more recent developments in hardware and software have been implemented by new departments. In some cases, in fact, implementation of advanced geocoding and GIS on microcomputers and workstations exceeded implementation of those systems on mainframes and minicomputers. This change can be explained as a result of the availability and low cost of the new computer systems of the 1980s. Prior to the 1980s, these systems were adopted only by departments with large budgets and jurisdictional sizes. For example, departments with mainframes and minicomputers use affordable and flexible mapping systems such as the U.S. Bureau of Census' Geographic Base Files, since these files are ready to be used and satisfy the planning departments' needs.

Organizational and institutional environments were discussed. Departments with mainframes and minicomputers use and maintain data in central units for planning and use by other city departments. The introduction of the 1960s and 1970s computers, and the experience gained from more than three decades of their use, provide the support to manage sophisticated and centralized systems. Also, it seems that these departments are surrounded by agencies with GIS technology. This finding indicates that the use and implementation of GIS technology are affected by many factors that have not yet been considered, such as types of government and the geographic location. This may be explained by the data-sharing capabilities of GIS and its subsequent desirability by other

agencies, state regulations, vendor locations, or marketing strategies. Research is needed to explain what major role these new factors play and how they play it in using GIS technology at the local government level.

As explained above, departments with GIS use and train many types of employees. The complexity of the early hardware/software systems required excellent skills to operate and ensure job security for a select number of employees. However, by the 1980s this situation was changing in such a way that less expertise was required to run microcomputer systems. Computing in the 1980s required more people, but with less experience and knowledge about computers. In addition, these new people could acquire experience within a few weeks or months and become experts in the use of workstations and microcomputer systems.

Finally, cluster analysis shows that planning departments were divided into five groups. The majority of them were classified as non-computer-oriented users, and a moderate number were classified as mainframe or new users. The last two groups were classified as CADD mapping departments and departments having large budgets and located in large-area cities. Geocoding and GIS technology have not been adopted extensively in local planning departments, where many departments have not used the technology yet. About ten percent of surveyed cities use geocoding on mainframe computers, another ten percent use GIS technology, and others use CADD and other systems. Departments' not having used geocoding and GIS technology in the past can be explained as a result of technological complexity and high costs. This phenomenon is changing so that some planning departments in small and large cities are adopting GIS technology and are conducting advanced planning applications more than those departments that use the old systems.

## **CHAPTER V**

### **ANALYTICAL RESULTS**

Local planning departments use advanced geocoding methods to provide the required, timely, and detailed information to planners, professionals, and decision makers. But does advancement in geocoding methods lead to improved data uses, increased spatial resolution, and increased volumes of data? Do the departments' environments play major roles in their use of geocoding technology? This chapter seeks answers to these two research questions.

Data for the variables of geocoding methods, data use, spatial resolution, data volume, and department budget were obtained from the questionnaire. Data for the variables of population and area were obtained from secondary sources. The geocoding methods, data use, and spatial resolution variables were measured at the ordinal level, while the data volume, area, population, and department budget variables were measured at the ratio level. Ratio variables were converted to the ordinal level because the purpose of this study is to examine how these variables are correlated. Variable operationalizing and data encoding are discussed in the first section. Since ordinal variables require special statistical analysis, nonparametric analysis is discussed in the second section. Research hypotheses are examined in the third section.

#### **VARIABLE OPERATIONALIZING AND DATA ENCODING**

Since obtained data are available at the ordinal and ratio levels of measurement, it is important to explain how they have been operationalized, grouped, and encoded. The

following sections show how ordinal variables (methods of geocoding, data use, and spatial resolution) have been operationalized and encoded, and how ratio variables (data volume, city area, city population, and department budget) have been operationalized, grouped, and encoded.

### Ordinal Variables

Methods of Geocoding. Six methods of geocoding were used to reference street addresses to geographic locations. They were:

- 1) Look up on map/look up in directory
- 2) Batch processing with tabular output
- 3) Interactive processing with tabular output
- 4) Transfer of geocoded data to a Thematic Mapping Program
- 5) Geocoding within a GIS
- 6) Geocoding within specialized systems

The seventh category, “Other”, was left open to be filled in by respondents so that they could explain any of their own methods of geocoding that were different from those listed. Respondents were asked to specify all methods of geocoding that applied. Very few designated the “Other” category only. Responses given in the “Other” category were matched with the appropriate geocoding methods provided in the question.

Answers to question 1 were used to measure the concept of advanced geocoding systems. The methods of geocoding were ranked from the least advanced (level 1) to the most advanced (level 6), with the highest-ranked answer for each case (266 cases) used to measure the level of advancement. Therefore, each case had only one answer. Table XXVII shows the types of geocoding methods and their levels of advancement. The frequency and percentage of each ranking are also provided.



TABLE XXVII  
METHODS OF GEOCODING AND THEIR RANKINGS FOR AUTOMATED  
AND MANUAL DEPARTMENTS

<u>Methods of Geocoding</u>	<u>Levels of Advancement</u>	<u>Frequencies N = 251*</u>	<u>Percentages N = 251*</u>
Look up on map/look up in directory	1	143	57%
Batch processing with tabular output	2	16	6%
Interactive processing with tabular output	3	20	8%
Transfer of geocoded data to a Thematic Mapping Program	4	6	3%
Geocoding within a GIS	5	53	21%
Geocoding within specialized systems	6	13	5%

N = Number of Departments

Source: Question 1

\* Missing cases = 15

Data Use. Question 2 asked respondents to specify all uses applicable to their departments. The uses were ranked from the least advanced (level 1) to the most advanced (level 4). The highest-ranked answer was used to measure the data use variable, as shown in Table XXVIII. Seventy-three departments did not specify any use, and omissions were considered missing data rather than assumed to mean “no use.” It

TABLE XXVIII  
DATA USES AND THEIR RANKINGS FOR AUTOMATED AND MANUAL DEPARTMENTS

<u>Types of Data Uses</u>	<u>Levels of Complexity</u>	<u>Frequencies N = 193</u>	<u>Percentages N = 193</u>
Tabulating addresses by statistical area without mapping the results	1	41	21%
Aggregating addresses to statistical areas and mapping	2	66	34%
Matching addresses to coordinates and/or network nodes	3	45	24%
Linking coordinate-based addresses to modeling and other applications programs	4	41	21%

N = Number of Departments

Source: Question 2

seems that about twenty-five percent of respondents do not use any of the listed uses. It is possible that they use geocoded addresses for other purposes but did not have the opportunity to report these uses.

Spatial Resolution. Six different levels of areal units were included in the question, and they were ranked from the largest to the smallest size. Spatially speaking, areal units were ranked from low resolution (1) to high resolution (6), and the highest-level answer was used to code the spatial data resolution variable. Only automated departments (112) were asked to specify all categories that applied. About one-fourth of the respondents specified additional areal units, but the sizes of these units varied and none were smaller than the smallest unit included in the question, i.e., ownership parcels. Ownership parcels were specified most frequently, followed by traffic zones. This phenomenon can be explained as a result of the use of land parcel data by a majority of planning departments. If land ownership data were geocoded to x,y coordinate spaces and their parcels pin-mapped, the spatial resolution would be correctly classified as geocoded to the land ownership level of resolution. However, many respondents specified the land ownership parcel category largely due to their geocoding of land parcel data to areas such as blocks, tracts, and planning areas. Consequently, the relationships between geocoding methods and spatial resolution, and between spatial resolution and data use and volume, were performed using five categories in which city blocks were considered the smallest geographic area to which the respondents geocoded data, as shown in Table XXIX.

TABLE XXIX  
SPATIAL RESOLUTIONS AND THEIR RANKINGS FOR AUTOMATED DEPARTMENTS

<u>Spatial Resolutions</u>	<u>Levels of Resolution</u>	<u>Frequencies</u> N = 81	<u>Percentages</u> N = 81
Planning areas	1	6	7%
Census tracts	2	8	10%
Traffic zones	3	14	17%
City block groups	4	8	10%
City blocks	5	45	56%

N = Number of Departments  
Source: Question 4

### Ratio Variables

The data volume, city area, city population, and department budget variables were converted to categorical variables. Appropriate grouping methods and the number of categories for each variable were determined based on the distribution of the scores on the line and the sample size. Monmonier (1981) explained four different methods of breaking ratio data. They are: 1) the natural breaks which separate clusters of points on the line, 2) quantile breaks which assign approximately the same number of areas to each cell, 3) equal breaks which divide the lowest and the highest values into equal parts, and 4) dispersion breaks which break objects with extremely high and low values. Based on the available data, dispersion breaks and four or five categories were the most appropriate groups and number of categories used for these ratio variables. The cutoff points and categories were determined by 1) estimating the mean and the standard deviation for each variable, 2) adding or subtracting 0.5 standard deviation to the mean, and 3) repeating step 2 until all elements in the sample were categorized.

Data Volume. Departments using automated geocoding methods were asked to specify all types of data that applied. Also, the quantity of transactions and frequency of updates were requested for each type of data. Data volume as a product of types and resolutions was impossible to compute due to missing and invalid information. Data

types or attributes should have been reported for each spatial resolution, but question 5 failed to yield such information. Instead, the volume of data transactions per year for each planning department was used to measure the data volume variable. This variable was calculated by adding all data transactions per year per case, and it was grouped into two categories, as shown in Table XXX, using the dispersion break method for automated departments.

TABLE XXX  
DATA VOLUME RANKING FOR AUTOMATED DEPARTMENTS

<u>Data Volume</u>	<u>Frequencies</u>	<u>Percentages</u>
	N = 55	N = 55
Less than or equal to 56,000	24	44%
More than 56,000	31	56%

N = Number of Departments

Source: Question 5

City Area, City Population, and Department Budget. As indicated earlier, city area and population were obtained from secondary sources. Planning departments' annual budgets were obtained from the questionnaire. Since these variables were measured at the ratio level, and the purpose of this study was to examine the correlation between methods of geocoding and these three variables, they were converted to the ordinal level. Eleven city areas and fifty-one department budgets were missing. Categories, ranks, frequencies, and percentages are provided in Table XXXI, using the dispersion break method and all responding cities.

TABLE XXXI  
CITY AREA, CITY POPULATION, AND DEPARTMENT BUDGET RANKINGS FOR AUTOMATED  
AND MANUAL DEPARTMENTS

<u>City Area in Square Miles (N = 255)</u>	<u>Ranks</u>	<u>Frequencies</u>	<u>Percentages</u>
Less than or equal to 20	1	58	21.8%
21-65	2	128	48.12%
66-110	3	37	13.91%
More than 110	4	33	12.41%
<u>City Population (N = 266)</u>			
Less than or equal to 170,000	1	192	72.18%
170,001-300,000	2	39	14.66%
300,001-430,000	3	17	6.39%
More than 430,000	4	18	6.77%
<u>Department Budget (N = 215)</u>			
Less than or equal to 500,000	1	92	34.59%
500,001-1,000,000	2	46	17.29%
1,000,001-1,500,000	3	32	12.03%
1,500,001-2,000,000	4	21	7.89%
More than 2,000,000	5	24	9.02%

N = Number of Departments

### NONPARAMETRIC STATISTICS

In order to measure the degree of association between two variables, the levels of measurement have to be identified as nominal, ordinal, interval, or ratio. Each of these scales uses a different type of statistical analysis to test associations among variables measured at one or more scales. Since ordinal measures were used to measure all variables in this study, nonparametric analysis was used to examine the degree of association between ordinal variables.

Nonparametric analysis is designed to measure the degree of association between two variables measured at the ordinal level. This type of analysis uses the extra information provided by the order of classes so that ordering leads to the concept of correlation. Nominal analysis, on the other hand, does not take advantage of such extra information. Positive correlation can be found when an object has low (high) values in the first variable and has low (high) values in the second variable. Negative correlation

shows the opposite tendency; high (low) values in one variable are correlated with low (high) values in the other variables. In the most general sense, correlation suggests only increasing or decreasing “monotonic” relationships, but they are not necessarily linear. Thus, ordinal measures of association are designed to measure how strongly two variables are correlated. For example, if two variables are exactly correlated, the coefficient will be +1.0; and if they are negatively correlated, the coefficient will be -1.0. If the coefficient is zero, it implies no correlation, and its absence does not mean statistical independence because the variables may be associated in another way (Reynolds, 1977).

Usually, correlation can be measured between two ranked variables. Since the independent variables are measured at the ordinal scale and the purpose of this study is to examine the correlation between ordinal and ratio variables, ratio variables have to be converted to ordinal variables. Although loss of information would be a problem in the converting process, it would not be serious if large amounts of data were used. The measure of correlation between two variables is the same as that between ratio variables. The difference is only arithmetic, but there are no new principles (Meddis, 1984).

Ordinal data were loaded into SYSTAT software and cross-tabulated, and the following nonparametric analyses were computed: Kendall’s tau-b, Stuart’s tau-c, Goodman and Kruskal’s gamma, and Spearman’s rho. These analyses are defined by the SYSTAT user manual as:

are appropriate when both categorical variables in a two-way table have ordered categories. The first three measures differ only in how ties are treated and the fourth is like the usual Pearson correlation except that the rank order of each value is used in the computation instead of the value itself (SYSTAT, 1989, p. 481).

A definition of each analysis is provided, followed by available methods used to test their significance as explained by Reynolds (1977), Kendall and Gibbons (1990), and

Meddis (1984). The appropriate analysis and test of significance were selected to test all research hypotheses.

#### Kendall's Tau-b

Kendall (1970, cited in Reynolds, 1977) first estimated the value of tau-a. It is at once the easiest measure to compute and understand, but the least helpful in practice. It measures the difference in probabilities that randomly drawn pairs are concordant or discordant. A concordant pair has two objects which have higher (or lower) values in both variables than the other object. A discordant pair has two objects; the first object has higher (or lower) values in the first variable but lower (or higher) values in the other variable. The problem with tau-a statistics is that tied pairs confuse their interpretation. Therefore, Kendall introduced a correction factor to remove the effect of ties and estimated a new coefficient called tau-b. The problem associated with tau-b is that it does not attain its maximum in nonsquare tables. Therefore, it is common to compute a different measure, tau-c.

#### Stuart's Tau-c

This statistic was developed for contingency tables and avoids some of the problems associated with tau-a and tau-b. It is designed to achieve its maximum in nonsquare tables. The only problem associated with this analysis is that it does not have a straightforward population interpretation. It is used mainly as a descriptive index to summarize the amount of correlation between two ordinal variables in rectangular tables.

#### Goodman and Kruskal's Gamma

This statistic provides the difference in probability of concordance and discordance among untied pairs. The problem associated with this statistic is that it is affected by the table size, since the number of tied pairs increases as the number of

categories declines; provided that the sample size remains constant, the absolute value of gamma increases. Conversely, as the number of categories increases, the absolute value of gamma decreases, all other things being equal (Reynolds, 1977).

### Spearman's Rho

This statistic computes the correlation between two ranked variables by subtracting two values for each pair, squares them, adds them, and divides them by the number of pairs. In the case of Kendall's tau-a and tau-b, each pair is assigned a value of +1.0 if any pair of objects are in the direct order and -1.0 in the reverse order; the statistic multiplies both results, adds them, and divides them by the number of pairs. Spearman's rho statistic can also be used for tied ranked data. A similar idea can be used to provide a measure of association in nonsquare contingency tables. The only difference between ranked and categorical data is that N refers to elements in the frequency tables, since each element refers to a pair of scores, and correlation looks for agreement between two ordinal variables rather than the difference between samples (Kendall and Gibbons, 1990; Meddis, 1984).

All of the above four analyses are applicable for ranked data as well as categorical data. Each one of these measures has advantages and disadvantages. Kendall's tau-b and Goodman and Kruskal's gamma are appropriate analyses for square tables, and are insensitive to tied ranks. Stuart's tau-c and Spearman's rho are appropriate measures for nonsquare tables, but tied ranks cause technical problems. Another problem associated with some of these techniques is the unavailability of the direct tests of significance. The following discussion shows how ordered variables can be tested for significance.

Kendall (1970, cited in Reynolds, 1977) presented a formula to test the significance of tau-b, assuming a large sample. It was designed to measure association among ranked data rather than among variables in contingency tables, although it could



also be used for the latter purpose. Goodman and Kruskal (1963, 1972) have given formulas for the standard error of gamma under various sampling conditions, but they are rather awkward.

Quade (1974) provided an asymptotic sampling theory covering most ordinal measurements of associations. He first defined a general index of correlation which found its asymptotic (e.g., large-sample) standard error. The formula was modified for contingency tables. Since Z values have an approximately normal distribution for large random samples, approximate two-sided confidence intervals can be obtained by multiplying the appropriate percentile of the unit distribution by the standard error. If the test value exceeds this result, the null hypotheses can be rejected.

This is an approximate estimate of the critical values, but it provides a fast and reasonably accurate method to test significance for contingency tables. In the case of Spearman's rho, Z and t values can be computed using available formulas and can be checked for significance in the available tables (Kendall and Gibbons, 1990; Reynolds, 1977; Meddis, 1984; Ferguson, 1981). Since Spearman's rho computation of the ties was not clear for contingency tables, Stuart's tau-c coefficient was used to test all research hypotheses and to test their significance.

## HYPOTHESIS TESTING

Eight hypotheses were examined in this study. The first five hypotheses examined the relationships between geocoding methods and data use, spatial resolution, and data volume, and between spatial resolution and data use and volume. The last three hypotheses examined the relationships between geocoding methods and city area, city population, and department budget. The relationships between geocoding methods and data use, city area, city population, and department budget were analyzed using 266

cases, while the relationships between geocoding methods and spatial resolution and data volume, and between spatial resolution and data use and data volume, were analyzed using only 112 cases because automated departments answered only spatial resolution and data volume questions (see questions 4 and 5).

Finally, Stuart's tau-c results were used to test the significance of all hypotheses, since it was the most appropriate measure for this study. Other statistics such as Kendall's tau-b and Spearman's rho were reported to show how these statistics were similar. The results of the statistical analysis for each relationship are presented and discussed in the following sections.

#### Methods of Geocoding and Data Use

The first hypothesis states that geocoding methods are significantly correlated with data use. Methods of geocoding (question 1) were cross-tabulated with uses of data (question 2), using categories presented in Tables XXVII and XXVIII. The results are presented in Table XXXII, which shows that advanced geocoding applications are significantly correlated with geocoding methods at the 0.01 level of confidence. Planning departments with advanced geocoding systems are capable of matching addresses to coordinates and/or networked nodes and of linking coordinate-based addresses to modeling and other application programs. When addresses are geocoded to coordinates, this linking enables local planners to integrate data from different sources; aggregate data by small geographic areas; and spatially analyze data for land use and transportation planning, facility locations, urban and housing development, and many other city planning applications. On the other hand, planning departments that use manual geocoding methods or limited automated systems can only tabulate addresses by statistical areas without mapping the results or aggregate addresses to statistical areas while mapping the results manually. As indicated earlier, about twenty-five percent of

both types of departments do not use address data. This is because they do not have the technology or they are not using address data in their departments.

TABLE XXXII  
THE RELATIONSHIP BETWEEN GEOCODING METHODS AND DATA USE  
FOR AUTOMATED AND MANUAL DEPARTMENTS

<u>Uses of Data</u>	<u>Methods of Geocoding</u>					
	<u>Manual</u>	<u>Batch</u>	<u>Interactive</u>	<u>Transfer</u>	<u>GIS</u>	<u>Specialized</u>
Tabulating addresses by statistical area without mapping the results	25	5	8	0	1	1
Aggregating addresses to statistical areas and mapping	39	7	8	3	9	0
Matching addresses to coordinates and/or network nodes	14	2	4	2	19	3
Linking coordinate-based addresses to modeling and other applications programs	9	0	0	1	22	9
Number of Departments	= 191	Standard error (SE)		Critical Value (1.96 x SE)		
Kendall tau-b	= 0.413	0.054		.106		
Stuart tau-c	= 0.395	0.053		.104*		
Spearman rho	= 0.481	0.059		.116		

\* = Significant at 0.01 Level of Confidence: (Critical Value < Stuart tau-c)

#### Methods of Geocoding and Spatial Resolution

The second hypothesis states that geocoding methods are significantly correlated with spatial resolution. Methods of geocoding (question 1) were cross-tabulated with spatial resolution (question 4) using categories in Tables XXVII and XXIX. The results are presented in Table XXXIII, which shows that spatial resolution is not significantly correlated with geocoding methods at the 0.05 level of confidence. Advanced geocoding systems are not the determinant for planning departments use of small geographic areas such as city blocks or ownership parcels for mapping and analysis. Departments with advanced geocoding technology are not necessarily using detailed data for planning and decision making. As a result, they use less-detailed information, which is often inappropriate for many planning applications at the city level.

TABLE XXXIII  
THE RELATIONSHIP BETWEEN METHODS OF GEOCODING AND SPATIAL  
RESOLUTION FOR AUTOMATED DEPARTMENTS

<u>Uses of Data</u>		<u>Methods of Automated Geocoding</u>			
<u>Spatial Resolution</u>	<u>Batch</u>	<u>Interactive</u>	<u>Transfer</u>	<u>GIS</u>	<u>Specialized</u>
Planning areas	1	0	1	3	1
Census tracts	1	1	1	4	1
Traffic zones	5	2	0	7	0
City block groups	0	0	0	5	3
City blocks	7	4	3	23	8
Number of Departments	= 81	Standard error (SE)		Critical Value (1.96 x SE)	
Kendall tau-b	= 0.073	0.093		.182	
Stuart tau-c	= 0.059	0.076		.149*	
Spearman rho	= 0.087	0.109		.214	

\* = Not Significant at 0.05 Level of Confidence: (Critical Value > Stuart tau-c)

#### Methods of Geocoding and Data Volume

The third hypothesis states that geocoding methods are significantly correlated with data volume. Methods of geocoding (question 1) were cross-tabulated with volume of transaction (question 5) using categories in Tables XXVII and XXX. The results are presented in Table XXXIV, which shows that there is no significant relationship between advanced geocoding systems and volume of data. Data volume may be affected by many other factors such as city area, city population, types of services, and types of applications. Also, the insignificant results were obtained because the majority of planning departments with advanced geocoding systems have small quantities of data transactions per year. Therefore, a correlation between geocoding and data volume cannot be found. However, the table shows that eight planning departments with GIS and specialized systems transact large quantities of data per year compared to four planning departments with batch, interactive, or transfer geocoded data systems. Advanced geocoding systems cannot be correlated with large volumes of data, especially when data volume is measured as the quantity of transactions per year.

TABLE XXXIV

THE RELATIONSHIP BETWEEN METHODS OF GEOCODING AND VOLUME  
OF DATA FOR AUTOMATED DEPARTMENTS

<u>Volume of Data</u>		<u>Methods of Automated Geocoding</u>				
		<u>Batch</u>	<u>Interactive</u>	<u>Transfer</u>	<u>GIS</u>	<u>Specialized</u>
Less than or equal to 56,000		8	7	2	18	8
More than 56,000		1	1	2	6	2
Number of Departments	= 55	Standard error (SE)		Critical Value (1.96 x SE)		
Kendall tau-b	= 0.063	0.112		.219		
Stuart tau-c	= 0.062	0.112		.219*		
Spearman rho	= 0.069	0.121		.237		

\* = Not Significant at 0.05 Level of Confidence: (Critical Value > Stuart tau-c)

### Spatial Resolution and Data Use

The fourth hypothesis states that spatial resolution is significantly correlated with data use. Spatial resolution (question 4) was cross-tabulated with data use (question 2) using categories in Tables XXIX and XXVIII. The results are presented in Table XXXV, which shows that spatial resolution is significantly correlated with data use at the 0.05

TABLE XXXV

THE RELATIONSHIP BETWEEN SPATIAL RESOLUTION AND  
DATA USE FOR AUTOMATED DEPARTMENTS

<u>Data Use</u>		<u>Spatial Resolution</u>				
		<u>PA</u>	<u>CT</u>	<u>TZ</u>	<u>CBG</u>	<u>CB</u>
Tabulating addresses by statistical area without mapping the results		1	0	1	1	3
Aggregating addresses to statistical areas and mapping the results		1	5	7	0	8
Matching addresses to coordinates and/or network nodes		3	3	3	0	11
Linking coordinate-based addresses to modeling and other programs		1	0	1	7	22
Number of Departments	= 78	Standard error (SE)		Critical Value (1.96 x SE)		
Kendall tau-b	= 0.259	0.084		.164		
Stuart tau-c	= 0.229	0.075		.147*		
Spearman rho	= 0.314	0.098		.192		

PA = Planning Areas

CT = Census Tracts

TZ = Traffic Zones

CBG = City Block Groups

CB = City Blocks

\* = Significant at 0.05 Level of Confidence: (Critical Value < Stuart tau-c)

level of confidence. Data geocoded by small geographic areas enable planning departments to conduct advanced geocoding applications such as matching addresses to coordinates and linking addresses for modeling and other programs. Data geocoded by large areas can be used for tabulation and aggregation, but not for matching and modeling. Small geographic areas can be flexible for data integration and modeling for planning applications.

### Spatial Resolution and Data Volume

The fifth hypothesis states that spatial resolution is significantly correlated with data volume. Spatial resolution (question 4) was cross-tabulated with volume of data (question 5) using categories in Tables XXIX and XXX. Table XXXVI shows that spatial resolution is significantly correlated with volume of data at the 0.05 level of confidence. Again, higher spatial resolution increases data volume so that many types of data attributes can be linked at smaller geographic areas. Addresses integrated by accurate geographic locations can be associated with data files; this integration increases the number of attributes for many geographic areas. Since data volume was measured by the quantity of transactions, geocoding by small geographic areas increases data volume.

TABLE XXXVI

#### THE RELATIONSHIP BETWEEN SPATIAL RESOLUTION AND DATA VOLUME FOR AUTOMATED DEPARTMENTS

<u>Data Volume</u>	<u>Spatial Resolution</u>				
	<u>PA</u>	<u>CT</u>	<u>TZ</u>	<u>CBG</u>	<u>CB</u>
Less than or equal to 56,000	5	4	8	3	19
More than 56,000	0	1	1	0	9

Number of Departments	= 50	Standard error (SE)	Critical Value (1.96 x SE)
Kendall tau-b	= 0.249	0.114	.223
Stuart tau-c	= 0.224	0.110	.215*
Spearman rho	= 0.261	0.120	.235

PA = Planning Areas

CT = Census Tracts

TZ = Traffic Zones

CBG = City Block Groups

CB = City Blocks

\* = Significant at 0.05 Level of Confidence: (Critical Value < Stuart tau-c)

Consequently, information about city blocks can be retrieved easily to answer questions. This significant relationship shows that small geographic areas provide up-to-date and detailed information required for planning applications. Data volume measured by attribute and resolution should also correlate with this finding.

#### Methods of Geocoding and City Area

The sixth hypothesis states that geocoding methods are significantly correlated with city area size. City area was cross-tabulated with methods of geocoding using categories in Tables XXXI and XXVII. The results are presented in Table XXXVII, which shows that geocoding methods are significantly correlated with city area size at the 0.05 level of confidence. Larger cities face the challenge of managing large quantities of spatial data. The only possible solution to processing large quantities of data is to use advanced geocoding systems in an effective and efficient manner. Hardware and software cost reductions and high memory capabilities encourage planning departments to handle their large geographic databases

TABLE XXXVII

THE RELATIONSHIP BETWEEN METHODS OF GEOCODING AND CITY AREA  
FOR AUTOMATED AND MANUAL DEPARTMENTS

<u>Area</u>	<u>Methods of Geocoding</u>					
	<u>Manual</u>	<u>Batch</u>	<u>Interactive</u>	<u>Transfer</u>	<u>GIS</u>	<u>Specialized</u>
Less than 20	36	4	6	1	6	2
21-65	74	6	9	0	27	7
66-110	14	5	1	2	7	3
More than 110	12	1	3	3	13	1
Number of Departments	= 243	Standard error (SE)		Critical Value (1.96 x SE)		
Kendall tau-b	= 0.171	0.052		.101		
Stuart tau-c	= 0.146	0.054		.105*		
Spearman rho	= 0.199	0.061		.109		

\* = Significant at 0.05 Level of Confidence: (Critical Value < Stuart tau-c)

automatically. As the table indicates, the majority of manual departments are in cities with small areas, less than or equal to sixty-five square miles, and an increasing number of small cities are using advanced geocoding and GIS.

#### Methods of Geocoding and City Population

The seventh hypothesis states that geocoding methods are significantly correlated with city population size. City populations were cross-tabulated with methods of geocoding using categories presented in Tables XXXI and XXVII. The results are presented in Table XXXVIII, which shows that geocoding and city population size are significantly correlated at the 0.05 level of confidence. Managing large volumes of data requires advanced geocoding systems to update and integrate data available among departments. In larger cities where most of the data are available in digital formats, planning departments are facing the challenge not only of automating the processing of their data but also of integrating data based on their geographic locations for the purpose of mapping and planning. Cities with large populations tend to have more advanced geocoding systems than cities with small populations.

TABLE XXXVIII

THE RELATIONSHIP BETWEEN GEOCODING METHODS AND CITY POPULATION  
FOR AUTOMATED AND MANUAL DEPARTMENTS

Population	Methods of Geocoding					
	Manual	Batch	Interactive	Transfer	GIS	Specialized
Less than or equal to 170,000	121	10	16	2	26	8
170,001-300,000	1	4	3	0	7	4
300,001-430,000	6	1	1	1	9	0
More than 430,000	2	1	0	3	11	1
Number of Departments	= 251		Standard error (SE)		Critical Value (1.96 x SE)	
Kendall tau-b	= 0.303		0.054		.105	
Stuart tau-c	= 0.211		0.040		.078*	
Spearman rho	= 0.339		0.060		.117	

\* = Significant at 0.05 Level of Confidence: (Critical Value < Stuart tau-c)



### Methods of Geocoding and Department Budget

The eighth hypothesis states that geocoding methods are significantly correlated with department budget. Planning department budget (question 15) was cross-tabulated with methods of geocoding using categories presented in Tables XXXI and XXVII. The results are presented in Table XXXIX, which shows that department budget is significantly correlated with geocoding methods at the 0.05 level of confidence. Although hardware and software costs are declining, departments with larger annual budgets are able to acquire and use advanced geocoding technology, but departments with smaller annual budgets still cannot afford it. This problem can be explained as a result of the high cost of spatial data capture, encoding, and formatting. Department budget still plays a major role in adopting and implementing advanced geocoding technology in the 1990s, but its effect is becoming weaker as costs decline.

TABLE XXXIX

THE RELATIONSHIP BETWEEN GEOCODING METHODS AND DEPARTMENT BUDGET  
FOR AUTOMATED AND MANUAL DEPARTMENTS

<u>Budget</u>	<u>Methods of Automated Geocoding</u>					
	<u>Manual</u>	<u>Batch</u>	<u>Interactive</u>	<u>Transfer</u>	<u>GIS</u>	<u>Specialized</u>
Less than or equal to 500,000	55	7	7	1	10	6
500,001-1,000,000	24	3	3	3	9	1
1,000,001-1,500,000	14	1	5	1	9	2
1,500,001-2,000,000	11	2	1	1	4	1
More than 2,000,000	11	1	1	0	10	0
Number of Departments	= 204	Standard error (SE)		Critical Value (1.96 x SE)		
Kendall tau-b	= 0.124	0.060		.117		
Stuart tau-c	= 0.105	0.051		.099*		
Spearman rho	= 0.1145	0.070		.137		
* = Significant at 0.05 Level of Confidence: (Critical Value < Stuart tau-c)						

### Summary

Advancements in geocoding systems enhance data use. Street addresses can be linked or modeled for mapping and analysis. Also, small geographic areas increase data

use and volume. Addresses geocoded by small geographic areas can be integrated with other data files, especially in the GIS environment, and provide detailed and up-to-date information. The correlations between geocoding methods and spatial resolution and data volume were found not to be significant, probably because spatial resolution was not measured properly and data volume was measured as the quantity of transactions per year, or because of other factors. City planning departments with advanced geocoding systems and GIS are located in cities with large areas, populations, and planning department annual budgets.

## CONCLUSIONS

The nonparametric analyses confirmed that most of the research hypotheses were correct. Advanced geocoding systems improve data use. Planning departments with advanced geocoding systems are able to conduct advanced geocoding applications. The capabilities of GIS technology have led city planning departments not only to map and aggregate data by large geographic areas, but also to match and link data to coordinates primarily needed for accurate mapping and data integration.

In terms of spatial resolution, the results show a significant correlation with data use and data volume. Even in departments with manual geocoding and less-sophisticated technology, small-area data help local planners to conduct advanced planning applications. Data can be retrieved easily for city blocks as long as they are identified and available. Detailed information can be used for mapping and spatial analysis using visual methods.

Large cities tend more than small cities to use advanced geocoding and GIS technology due to their experience with computer hardware, software programs, and database structure and maintenance. The availability of local professionals and large

annual budgets have encouraged planning and development directors to adopt this technology for a long time. Technological advancements are eliminating organizational and institutional problems, but creating new difficulties. This analysis showed that planning department budget is weakly correlated with geocoding methods, but city area is strongly correlated.

Only two hypotheses were not confirmed. Advanced geocoding systems do not enhance spatial resolution nor do they increase data volume. This result could be related to the measurements in which variation in spatial resolution was not possible and the quantity of transactions was used instead of the quantity of attributes and resolution. However, many planning departments with advanced geocoding systems use small geographic areas and transact large quantities of data per year.

In short, geocoding technological advancements have positive impacts on data use. Spatial resolution can improve street address utilization and provide up-to-date information. The adoption and implementation of geocoding technology can be related to organizational, institutional, and technical environments. However, some of these environments are changing as technology becomes more advanced.

## CHAPTER VI

### DISCUSSION AND CONCLUSIONS

This study examined the relationships between technological advancements in geocoding systems and data use, spatial resolution, and data volume; between spatial resolution and data use and volume; and between technological advancements in geocoding systems and area, population, and department budget. The results were new and important for GIS design and implementation. Usually, vendors and users make decisions to design and use such technology without knowledge about technical needs and sometimes with little knowledge about organizational and institutional environments. Information from this study, and hopefully others related to it, will help fill these voids.

The findings of this study explain how technical, organizational, and institutional environments are related, in order to facilitate more effective use of GIS in city planning departments. These ideas are considered further beginning with a discussion of the generalizability of the study, followed by implications for GIS theories and acquisition, and ending with topics which need further research.

This study surveys planning departments in cities with populations of 50,000 or more in the United States. As noted earlier, responding cities have area sizes and populations similar to those of non-responding cities (see Appendix B, Table B-I). Findings of this study can be generalized to apply to all American cities with populations of 50,000 or more in the U.S. Also, these findings can be generalized to apply to towns and boroughs with similar area sizes and populations. The findings of this study may not be generalized to apply to other departments of the same size, cities with populations of

less than 50,000, and other local government agencies, since spatial resolution and data volume are very sensitive issues in designing geocoding and GIS technology.

### THEORETICAL IMPLICATIONS

Street address data in local planning departments are used for many planning applications. Addresses have to be geographically located for the purposes of tabulating by statistical areas, mapping to show location of objects, using an address as a location to integrate all related data to that address, or linking data for modeling and programs using location in the modeling process. For one or all of the above to be accomplished, addresses have to be referenced (geocoded) to specific geographic locations using one or more of the available geocoding methods. The first and the most common method of geocoding addresses is the manual process, by which addresses are located on maps or addresses are assigned by geographic codes such as census tract number or area identification number. Since the computer revolution about four decades ago, the geocoding process has been conducted automatically. Furthermore, the automated process has become more advanced using available hardware systems and software programs and accurate map bases available in digital forms at reasonable costs.

The geocoding literature has emphasized the issue of advancements by which accurate and efficient processing of geocoding can be performed. For example, addresses can be assigned to accurate locations as a result of the use of accurate maps and sophisticated matching software programs. With the availability of relational database structures, addresses can be manipulated with data records if they are assigned by identification numbers. Geocoding researchers believe that such advancements should have positive impacts on data use, spatial resolution, and data volume using street addresses as the location identifiers. Other researchers have found that organizational and

institutional environments play major roles in the adoption and implementation of geocoding technology. A detailed discussion of how technology and organizational and institutional environments are in agreement with theoretical thoughts is needed. The results of this study help to answer this question.

### Technical Issues

Manual geocoding is an inaccurate and time-consuming process. If there are large quantities of address data files that must be merged and analyzed on a day-to-day basis, locating these addresses on maps is impossible. In addition, if addresses having more than one attribute and are describing more than one purpose, manual geocoding, mapping, and tabulation would be impossible. In an automated environment, where addresses and their attributes are stored, manipulated, and analyzed with the use of computer systems, addresses can be used more frequently, and changes can be made more quickly. Also, addresses can be mapped in an efficient and effective manner. The role of computer systems is the same as the role of humans, except that computer systems process data much more quickly and accurately than humans.

When interaction with technology is possible, people can monitor mistakes and changes during the operation rather than waiting until processing ends. More work can be done and new information can be added or dropped before the job is finished. The evolution of interactive software programs was a positive step toward solving the problems of batch software programs. The batch and interactive geocoding process was done on mainframe computers using software written in the programming language FORTRAN, and was conducted by professional personnel, usually systems analysts and computer programmers.

By the time hardware became smaller, software programs became more common and user-friendly, employees became more aware of computer technology, and

geographic databases became more available in digital format, so that addresses could be geocoded, manipulated, and mapped quickly using available and very simple technology. Addresses can be associated with geographic areas using polygon identifiers to perform thematic mapping products. Different attributes can be mapped showing distribution of employees, income, crimes, and land use. The last few years have seen an increase in the availability of compatible hardware, user-friendly software, and more importantly, accurate and comprehensive map bases for each part of the nation. Directors, managers, and decision makers can now geocode addresses with locations, integrate data from different sources, and map data within a few hours. This study showed that advancements in geocoding technology enable planning professionals to handle their problems without depending on computer professionals. Furthermore, planning and development directors and their staff are able to create new methods of integrating and mapping data in more efficient and effective ways to study different alternatives and approaches.

Also, this study showed that planning departments with GIS use small area data, which helps in conducting advanced planning applications and generates more data for planning and management. Working with advanced geocoding systems ensures that addresses are assigned unique identifier numbers or x,y coordinate locations. Processing data based on their locations enhances data use and increases the flexibility, efficiency, and accuracy of data integration, particularly the important step -- aggregation to small area data. Also, data retrieval and updating can be done more easily and quickly.

Although geocoding literature suggests that advanced geocoding methods increase spatial resolution and data volume, this study did not confirm these hypotheses. This could be because planning departments do not need to integrate many attributes by small

geographic areas, or because measuring data volume as the quantity of transactions per year was not a valid method of measurement.

Finally, this study showed that the adoption of technology has moved from mainframe computers to minicomputers to microcomputers to workstations. Software programs have moved from office automation programs to CADD to GIS programs. Advancements in geocoding and GIS technology led to better data use and spatial resolution, and up-to-date information required for mapping and planning applications is available at the local level. Because of the high cost of abandoning investments in mainframe computers, departments with these investments are slower to invest further in modern technology. Conversely, departments without these investments can readily procure the new technology available on microcomputers and workstations.

#### Organizational and Institutional Issues

The theories of centralized and decentralized systems are commonly used in defining government organizations. Since geocoding and GIS technologies are multidisciplinary tools, neither centralized nor decentralized systems are applicable. Another theory is called networking, in which all computers in an organization are located in all departments and are linked. Each department has its own hardware, software, and database, and all departments can share and use all databases throughout the network. However, this study showed that decentralized units were a significant factor in separating departments into the five groups, as explained earlier. This means that city organizations have not changed to meet the new technology.

The geocoding literature suggests that organizational and institutional environments determine how advanced local governments are in adopting and implementing geocoding technology. Previous studies indicated that cities that have large populations, annual budgets, and numbers of full-time employees, and strong forms



of government (city managers), are more likely to adopt and implement geocoding technology. These findings confirm those of previous research indicating that cities with large populations and annual budgets use advanced geocoding and GIS technology more than cities with small populations and annual budgets, although technology has advanced and changed.

The analysis showed that some of the institutional variables are becoming very weak; cluster analysis indicated that department budget was the weakest variable, and it was not significant. City population was ranked the fifth-strongest variable. Thus, new institutional variables have emerged and play major roles in the adoption and implementation of technology. For example, city area was the strongest variable segregating clusters. What is not known is how low- and high-density cities, slow and fast growing, new and old, are investing in GIS, and investigation in this area is needed.

As discussed in the previous section, city organization and awareness by departments of advanced geocoding and GIS technology have also had major impacts on the adoption and implementation of technology. It seems that old obstacles and barriers were removed as a result of technological advancements, but new problems and difficulties are arising. How new problems and difficulties can be solved is the question addressed in the following section.

## IMPLICATIONS FOR DEVELOPMENT AND ACQUISITION

Local planning departments use street address data for many planning applications. The manual method of geocoding is not going to last forever, and the demand is increasing to process, update, and provide detailed data for planning, management, and decision making. Current results showed that a large number of local planning departments are acquiring and using GIS. Also, many respondents who do not

use GIS have said that they are in the process of considering and obtaining such technology. So far, methods and strategies have been drawn for users and vendors on how to get and use GIS. Most of the available strategies and approaches are available in current GIS literature; each of these sources has designated one chapter or section for GIS implementation. Most of these methods and strategies were adopted from general theories of planning and decision making, and they have been modified and updated for the adoption and implementation of GIS technology.

As indicated in the previous section, GIS technology is a multidisciplinary field in which technical, organizational, and institutional environments are interrelated. In order for one of the above implementation approaches to be used, all of these environments have to be considered simultaneously. In the following sections, first, technical issues are discussed; second, organizational and institutional issues are addressed.

### Technical Issues

Hardware, software programs, and databases are usually the technical issues considered before a GIS is acquired, and only after objectives and required data have been determined. As shown in this study, types of applications, spatial resolution, data volume, and methods of geocoding have to be discussed as early as possible before the implementation plan starts. Even if objectives and data have been clearly defined, hardware and software programs have been carefully selected, and spatial and attribute data have been perfectly structured, the use of advanced geocoding and GIS technologies will still depend on how data and addresses have been geocoded, what advanced applications can be created, how and why small geographic areas are used, and how large the database is. Determining each of these issues is an important step in the implementation plan.

The issue of geocoding methods is an important one to be identified in the implementation process, because any additional information linkage depends on how data records are structured and which identifier is used for geocoding. For example, with address data, an address is assigned to a geographic location by tagging a location identifier such as city block number (in nominal geocoding) or x,y coordinate value (in positional geocoding). The process of addressing data for tabulation, mapping, and integration depends on which identifier is used. If city block number is used, data associated with street addresses can be related to city blocks and can be integrated with other data tabulated by city block.

On the other hand, if positional geocodes have been used and assigned to address data records, address data can be linked with any geographic area around that coordinate location, and data can be tabulated, mapped, and integrated with other data files if they are related to the same areal unit. In advanced geocoding systems, the process of positional geocodes can be related to any geographic location in the form of points, lines, and areas. In this case addresses can be managed and analyzed using their coordinate locations rather than identification numbers.

The other issue that should be considered in the implementation process is spatial resolution. Two problems are associated with this issue: first, the size of areal units, and second, data integration. If small geographic areas are selected and a large amount of data will be generated, detailed information will be available and data can be aggregated by larger geographic areas without any loss of any information. On the other hand, if large geographic units are selected and a small amount of data is produced, little information will be available. Also, disaggregating large units may cause some information to be lost, unless advanced geocoding systems are used to avoid such problems. The sizes of area units have to be determined before hardware and software

programs are selected and spatial database structuring begins. Then an appropriate level of spatial resolution will be used without causing problems in analysis, integration, and memory space.

The last issue that must be considered in the implementation stage is determining database size. Since data volume is a product of resolution and data types, data required for planning applications identified by the objects and memory space have to be determined for the database. Large volume of data are difficult to measure. Thus, large amounts of unnecessary data are undesirable. Depending on the department objectives, types of applications, types of data, and spatial resolutions, data volume can be determined and memory space can be estimated. Other technical issues such as data quality, hardware and software programs, and spatial data structure have to be considered carefully before the implementation plan starts.

#### Organizational and Institutional Issues

Technology cannot work by itself. Personnel are needed to enable technology to work. Personnel, on the other hand, need funding and organizational structures to perform the job. Therefore, personnel, budget, and administration have to be considered as part of the implementation plan. The geocoding literature stresses these concerns, and they should be addressed as part of the planning process. Centralized and decentralized approaches were the only suggestions for city organization. Either approach would be appropriate, depending on the organization's objectives and structure. In GIS environments, where data are used by many departments in an organization, a networking approach is appropriate to avoid the disadvantages of the centralized and decentralized approaches. All departments share the database, and each department has its own attribute information. If changes occur at the map base, all departments get the change; if the changes occur in one department (such as adding more data attributes to ownership

parcel), planning departments can keep the new changes in their databases and not affect other department databases. This approach is useful, since advanced geocoding systems require detailed and accurate map bases, and attribute data can be shared and maintained by each department. Some of the returned questionnaires indicated that the responding departments share map bases and each department has its own data attributes.

This study has shown that other organizations in the metropolitan area use GIS, indicating they might use the same maps as planning departments or use other maps in the area. Networking can be expanded to other organizations to enhance and standardize mapping operations. For example, if an agency has updated and corrected TIGER files or developed its own maps, other organizations in the area can use these maps, and they do not have to worry about creating or editing new maps. Using available maps saves money and time, and standardizes data.

In terms of the institutional environment, large-area cities with large populations and department budgets are more likely to use geocoding and GIS technologies. These issues are still affecting development and acquisition, but these conditions are becoming weaker, especially for the department budget variable. Adoption and implementation of GIS technology is no longer dominated by larger cities, because GIS are becoming available and affordable for small communities as well. Many respondents have indicated that they are considering GIS at this time. Attention has to be paid to new institutional variables as a result of technological advancements, not only to overcoming old obstacles and barriers. The challenge that is going to face local governments, especially large ones, is how to solve old problems without creating new ones.

Finally, in order to adopt and implement an advanced geocoding system, technical, organizational, and institutional environments have to be considered as early as possible in the implementation plan. All environmental issues have to be considered at

once to solve old problems and anticipate future problems and difficulties. This study shows that as technology is becoming more advanced, organizational and institutional environments are having fewer problems, but new problems should be expected. For example, hardware, software programs, map bases, employees, and training are available at reasonable costs. But which hardware and software programs are suitable for planning applications, what types of employees are needed, and what type of training is required? Planning departments have to know in advance what their requirements and needs are in order to make accurate implementation plans which consider all of the above issues rather than considering each one separately.

### RESEARCH NEEDS

This study shows the importance of research. However, more studies on GIS technology and geographic data are needed. Such studies would determine whether the results of this study are unique to city planning departments or generally applicable. Two areas of research are needed: technical research and organizational and institutional research.

#### Technical Research

This study has raised some questions about spatial resolution and data volume, which were not addressed in this study and need further research. First, are ownership parcels used for planning, and if not, why not? In other words, do local planning departments use the highest spatial resolution for planning? This information is important for system design. Second, do planning departments geocode data to segments and points? Both of these issues are important for determining spatial resolution and data volume needed in system design.

The study shows that planning departments use Level I spatial resolution, while public works and engineering departments use Levels II and III. Studying these departments may give detailed information about resolution and data volume, which would help to determine these relationships. Case studies or interviews are recommended for collecting this type of detailed data.

#### Organizational and Institutional Research

Departments were examined in terms of geocoding methods and data use, spatial resolution, data volume, area, population, and department budget. One area of research could address the question of why non-computer-oriented and CADD mapping departments do not use geocoding and GIS technology. Another area of research could address the fact that technological development is increasing at an exponential rate requiring up-to-date information on organizational and institutional environments. This study shows that many organizational and institutional elements are in flux. Research on organizational and institutional environments is necessary to determine how GIS is related to these new environments.

Another type of research is needed to compare different organizations adopting this technology at the local, state, and federal levels. This kind of research could compare institutional environments in different settings and could help others to adopt this technology. Local governments consist of many departments that are interdependent. For example, planning departments must consider water supply, sewer systems, roads, utilities, and police and fire protection. Further research directed toward all departments would strengthen the relationships among these different departments. In addition, this interdependency can be seen to extend beyond the local level to the relationships of cities to the state and states to the country as a whole.

Finally, due to affordable and user-friendly technology, small communities (e.g., cities with populations under 50,000) are turning to GIS for planning and many other purposes. Research is needed to determine their requirements and the types and sizes of systems they have, because they have limited resources and different environments.

## CONCLUSIONS

Many planning departments use advanced geocoding and GIS technology to provide the up-to-date and detailed information required for planning, management, and decision making. Many organizations are using advanced geocoding methods which help to tabulate, aggregate, match, and model address data needed for many planning applications. Fine spatial resolution has been used for geocoding and planning, which provides detailed information that was not available before. Both large and small cities are adopting advanced geocoding and GIS technology because costs have decreased and systems have become more user-friendly. Despite technological advancements and the removal of organizational and institutional barriers and obstacles, new problems and difficulties are arising. The challenge is how to determine and predict such problems. The solution is to know as much as possible in advance and create a comprehensive plan, rather than taking the plan step by step and solving problems as they occur. Considering technical, organizational, and institutional environments simultaneously for GIS development and acquisition is necessary and is the research topic of GIS for the 1990s.



## SELECTED BIBLIOGRAPHY

- Aangeenburg, Robert T. and Dorothy L. Bomberger (editors). 1975. International Geocoding '74. Colloquium Proceedings. Second International Colloquium on Segment-Oriented Referencing Systems, Wichita, KS, Institute for Social and Environmental Studies. The University of Kansas.
- Al-Ankari, K. M. 1991. "An Incremental Approach For Establishing a Geographic Information Systems In A Developing Country: Saudi Arabia," International Journal of Geographic Information Systems, Taylor and Francis, Washington, DC, Vol. 5, No. 1, PP. 85-98.
- Alexander, David. 1989. "GIS-Making It Low on Cost and High on Benefits," In Proceedings of the Ninth Annual Environmental System Research Institute (ESRI) Users Conference, Palm Springs, CA, 1989, May 22-26.
- American Congress of Survey and Mapping (ACSM). 1989, August. "Multi-purpose Geographic Data Base Guidelines for Local Governments," Washington, DC., pp. 42-50.
- The American Farmland Trust. 1985. "Survey of Geographic Information Systems for Natural Resource Decision Making at the Local Level." Washington, DC.
- Appleton, Daniel S., Kenneth L. Kraemer and David G. Schetter. 1977. A National Urban Information System Resource Center: Assessment of Need and Concept, The Public Policy Research Organization, University of California, Irvine, CA.
- Arbeit, David, Gerry Clancy and Mary F. Martinez. 1988. "Feeding a Data Hungry Public: Creative Uses of New Computer Technology," URISA, Vol. IV, pp. 330-331.
- Aronoff, Stan. 1991. Geographic Information Systems: A Management Perspective, WDL Publications, Ottawa, Canada.
- Bainbridge, William. 1989. Survey Research: A Computer Assisted Introduction, Wodsworth Publishing Company, Belmont, CA.
- Barabba, Vincent P. 1974. "Statistical Data," In the U.S. Bureau of the Census 1974. Geographic Base Files: System Users, Maintenance, Problem Solving, Conference processing, Series GE 60, No. 3, Arlington, Texas. Washington, DC: U.S. Bureau of the Census, pp. 65-69.
- Barb, Charles Elliot, Jr. 1971. "Rocky Road to 1980," Geocoding 71, Cambridge, MA: URISA.

- Barb, Charles Elliot, Jr. 1974. "Automated Street Address Geocoding Systems: Their Local Adaptation and Institutionalization," (unpublished dissertation), The University of Washington, Seattle, Washington.
- Barnell, Brad. 1989. "Using Problem Matching and Photogrammetry to Create Accurate Parcel Data," Proceedings, ESRI.
- Becker, Patricia . 1980. "Using Census Data to Develop Planning Analysis," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Bishton, Amy. 1988. "Designing and Using a Cartographic Extract: Mapping From the TIGER System," Proceedings, 1988 Annual Conference of the Urban and Regional Information Systems Association (URISA), pp. 130-141.
- Brooks, Kristina. 1985. "Determination of Geographic Information System Database Characteristics: An Application of the Information Integrated Theory," Doctoral Dissertation, Oregon State University, Corvallis, OR.
- Bryan, Nora. 1989. "Computerized Street Files, Still Useful and Versatile: Second Case Study Applications," URISA, Vol. IV, pp. 106-114.
- Bryan, Nora Sherwood, Donald F. Cooke, James Guthrie, Deborah A Reed and Joel Z. Yan. 1991. "Taking Stock: The Applicability of Street Centerline Files to GIS in the 1990s." URISA, Vol. II, pp. 198-206.
- Bureau of the Census, 1990. 1990 Census Population, Washington, D C: U.S. Government Printing Office.
- Burrough, P. A. 1987. Principles of Geographic Information Systems for Land Resource Assessment, New York, NY: Oxford University Press.
- Calkins, Hugh. 1972. "An Information System and Monitoring Framework for Plan Implementation and the Continuing Planning Process," Doctoral Dissertation, University of Washington, Seattle, WA.
- Cater, Erlet. 1974. "Computerized Information Systems for Planners," in Jean Perraton and Richard Baxter (eds.), Models, Evaluation and Information Systems for Planners, Land Use and Built from Studies (LUBFS), Conference Processing No. 1, University of Cambridge, Department of Architecture.
- Chamberlain, Simon B. 1980. "Land Use Planning: Toronto Case Study," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Chrisman, Nicholas R. and Bernard J. Niemann, Jr. 1985. "Alternative Routes to a Multipurpose Location: Merging Institutional and Technical Reasoning," Auto-Carto 7, Proceedings: Digital Representation of Spatial Knowledge, American Society of Photogrammetry and American Congress of Survey and Mapping, pp. 84-94.

- Clark, William L. 1967, April. "Urban Geocoding Systems and Their Utility," Threshold of Planning Information Systems: Selected Papers Presented at the ADP Workshop Conducted at the ASPO National Planning Conference, Houston, April 1967, pp. 57-59.
- Clarke, Keith C. Analytical and Computer Cartography, Chapter IV Geocoding. Prentice-Hall, Inc. NJ.
- Cook, Thomas, 1989. "Choosing Mapping Software," American City and County, Vol. 104, No. 6, p. 14.
- Cooke, Donald F. 1980. "A Review of Geoprocessing Systems and a Look at Their Future," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Corbett, James and George Farnsworth. 1976. "Theoretical Basis for Dual Independent Map Encoding (DIME)," Geocoding 71, pp. 22-31.
- Corbett, James P. 1973. "The Use of the DIME System for Identification of Places and Description of Locations," in Moyer and Fisher (eds.), Land Parcel Identification for Information Systems, pp. 1-149 - 1-155.
- Dangermond, Jack, 1989. "A High-Tech Alphabet Soup," American City and County, Vol. 104, No. 7, p. 12.
- Dangermond, J. and C. Freedman. 1986, "Findings Regarding a Conceptual Model of a Municipal Data Base, and Implication for Software Design," Proceedings to the International Symposium on Spatial Data Handling, Zurich, University of Zurich-Irchel, pp. 479-496.
- Dangermond, Jack. 1989, "A Classification of Software Components Used in Geographic Information Systems," Basic Reading in Geographic Information Systems, edited by Donna Peuquet and John O'Callegan. 1984 (Williamsville, NY, SPAD, System Ltd).
- Danziger, J. N. and Kenneth Kraemer. 1986. People and Computers: The Impacts of Computing on End Users in Organizations, New York, NY: Columbia University Press.
- Darnell, Tim and Jennifer Carlile, 1989. "Users Blues," American City and County, Vol. 104, No. 7, pp. 36-41.
- Darnell, Tim. 1988. "Applying High-Tech to Local Operations," American City and County, Vol. 103, No. 7, pp. 34-34.
- Dillman, Don A. 1978. Mail and Telephone Surveys: The Total Design Methods, New York, NY: John Wiley and Son.
- Dueker, K. J. 1988. "Urban Applications of Geographic Information Systems Technology: A Grouping Into Three Levels of Resolution," URISA, pp. 104-109.

- Dueker, K. J. 1987a. "Geographic Information Systems and Computer Aided Mapping," Journal of the American Planning Association. Vol. 53. pp. 383 - 390.
- Dueker, K. J. 1987b. "Multi-purpose Land Information Systems: Technical, Economic, and Institutional Issues," Photogrammetric Engineering and Remote Sensing, Vol. 53, No. 10. pp. 1361-1365.
- Dueker, K. J. 1980. "An Approach to Integrated Information Systems for Planning," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Dueker, K. J. 1979. "Land Resource Information Systems: A Review of Fifteen Years Experience," Geo-Processing, 1(2), pp. 105-28.
- Dueker, K. J. and Daniel Kjerne. 1989. Multipurpose Cadastre: Terms and Definitions, American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, Falls Church, VA.
- Durfee, R. C. 1988, February. "Review of Geoprocessing Technology Applicable to Regional Analysis," Paper prepared for the Environmental Protection Agency for the U.S. Department of Energy.
- Environmental Science Research Institute (ESRI). 1986. ARC/INFO Users Manual, Redland, CA.
- Epstein, Earl F. 1989. "The Impact of Information Science on Conflict," URISA Vol. IV, pp. 305-309.,
- Eveland, J. D. 1980. "Applying GBF/DIME and Other Geocoding Systems," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- French, Steven P., Lyna L. Wiggins and Kelly M. Hellernon. 1989. "California Planning Agency Experience With Automated Programming and Geographic Information Systems," Paper presented to the American Collegiate School of Planning 31st Annual Conference, October 5-7, 1989, Portland, OR.
- French, Steven P. and Heather L. McMillan. 1989, March. "San Luis Obispo County User Needs Assessment for Automated Mapping/Geographic Information Systems." California Polytechnic State University. San Luis Obispo, CA.
- Geocoding-71, Papers from the Working Session on Geographic Base File Development, 9th Annual Conference of the URISA September 8-10, 1971, New Orleans, Louisiana, URISA.
- Gordon, Kenneth E. 1981. "Environmental Data Use in Computer Assisted Spatial Data Handling Systems: The Results of a Survey of Applications in the Pacific Northwest States," Paper No. 15, Center for Pacific Northwest Studies, Washington State University.

- Harris, Britton. 1989. "Beyond Geographic Information Systems: Computers and the Planning Professional," Journal of the American Planning Association, Vol. 55, No. 1, pp. 85-90.
- Holtzheimer, Paul Edgar. 1983. "An Introduction to the GBF/DIME: A Primer," Computer, Environment and Urban Systems, Vol. 8, pp. 133-173.
- Horwood, Edgar M. 1967. "State of the Art of Planning Information Systems," In Threshold of Planning Information Systems, selected papers presented at the ADP Workshop conducted at the ASPO National Planning Conference, Houston, TX, pp. 1-3.
- Horwood, Edgar M. 1980. "Planning Information Systems: Functional Approaches, Evaluation, and Pitfalls," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Huxhold, William. 1980. "Planning Agencies and the Data Planning Department: Working Together Effectively," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Huxhold, William E. 1991. An Introduction to Urban Geographic Information Systems, Oxford University Press, NY.
- Hysom, John and Stephen Ruth. 1984. "A Nationwide Assessment of Local Government Planning Information Systems," Computer, Environment and Urban Systems, Vol. 9, pp. 117-126.
- Innes, Judith. 1988. "Effects of Data Requirements on Planning: Case Studies of Environmental Impact Assessment and Community Development Block Grants," Computer, Environment and Urban Systems, Vol. 12, No. 2, pp. 77-88.
- Jin, Runghua and Daryl Smith. 1989. "A Pilot Study on Relating Records to Cover Features by Structure Pattern Matching," ESRI, Redland, CA.
- Johnson, Mary Jean. 1989. "Mapping the Way: Geographic Information Systems Showing Payoff," Computer World, Vol. 23, p. 23.
- Juhl, Ginger. 1989, September. "Cincinnati's Strategy for Regional GIS," Consumer Technology, Vol. 2, No. 5, pp. 25-27.
- Kendall, Maurice and Jean Diekinson Gibbons. 1990. Rank Correlation Methods, Oxford University Press.
- Kindleberger, Charles P. 1988. "Planning Support System for the 1990s: Local Government Information Processing Challenges and Opportunities," URISA, Vol. III, pp. 1-21.
- Kinnear, Christine. 1987. "The TIGER Structure," Auto-Carto 8, Proceedings, American Society of Photogrammetry and American Congress of Survey and Mapping, pp. 249-257.

- Kinzy, Steven W. 1977. Geoprocessing in Local Government: Feasibility Study for the City of Omaha, URISA, Professional Paper Series, Institute of Urban and Regional Research, University of Iowa, Iowa City, IA.
- Kinzy, Steven W. 1980. "Urban Geoprocessing: A Staged Development Concept." Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Klosterman, Richard E. 1987. "Guidelines for Future Computer-Aided Planning Models," URISA, Vol. IV, pp. 1-14.
- Kraemer, K. L., W. H. Mitchell, M.E. Weiner, and D. E. Dial. 1974, Integrated Municipal Information Systems: The Use of Computers in Local Governments, Praeger Publication, NY.
- Kraemer, K. L., W. H. Dutton and Alan Northrop. 1981. The Management of Information Systems, New York, NY: Columbia University Press.
- Kraemer, K. L. and John Leslie King. 1977a. Computers and Local Government: A Manager's Guide, Vol I, New York, NY: Praeger Publishers.
- Kraemer, K. L. and John Leslie King. 1977b. Computers and Local Government: A Review of Research, Vol. II, New York, NY: Praeger Publishers.
- Kraemer, Kenneth and Mitch Modeleski. 1975. "Geosites: An Empirical Study of Geoprocessing in U. S. Cities and Counties," In U.S. Bureau of the Census (1975) Computerized Geographic Coding Series GE 60, No. 7: A Geographic Dimension for Decision-making, Washington, DC: U.S. Bureau of the Census, pp. 14-25.
- Larsen, B. et al. 1978. "Land Records: The Cost to the Citizen to Maintain the Present Land Information Base," a case study of Wisconsin's Department of Administration, Madison, WI: State of Wisconsin.
- Levine, Johnathan and John D. Landis. 1989. "Geographic Information Systems for Local Planning," Journal of the American Planning Association, Vol. 55, No. 2, pp. 209-220.
- MacGaffey, Neil and Hava Shalit, 1988, "Address Geocoding in GIS Development: Technical Considerations From a Case Study," GIS/LIS, pp. 190-197.
- Marshall, Martha W. 1988. "A Local Government's Approach to Developing an Integrated Land Information System: Prince William County's Experience," URISA, Vol. IV, pp. 195-206.
- Marx, R. 1986a. "Computer Mapping and The Census Bureau: The Future and You," Paper presented to the AM/FM International Eastern Regional Conference, Williamsburg, Virginia, May 5-7, 1986.
- Marx, R. 1986b. "The TIGER System: Automating the Geographic Structures of the United States Census," Government Publications Review, Vol. 13, pp. 181-201.

- Marx, Robert W. 1987. "The TIGER System: Six Years to Success," Paper presented at the 13th International Conference of the International Cartographic Association, Moralia, Mexico, October 1987.
- Marx, Robert W. 1989. "GIS, TIGER, and Other Useful Acronyms," Paper presented at the National Conference of GIS, Ottawa, Ontario, Canada.
- McGregor, Darrell. 1988. "GIS Not Just for Mapping," American City and County, Vol. 103, No. 12, p. 10.
- Meddis, Ray, 1984. Statistics Using Ranks, Basil Blackwell Inc., NY.
- Monmonier, Mark S. 1981. Maps, Distortion, and Meaning, Resource Paper No. 75-4, Association of American Geographers, Washington, DC.
- Metzgar, Robert. 1980. "Geoprocessing Systems: City of Long Beach Case Study," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Morris, Joe. 1987. "Computer Mapping the Infrastructure," American City and County, Vol. 103, No. 4, pp. 50-58.
- Moyer, D. David. 1980. "Land Records Reform and the Future of the Land Records Systems," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ Auerbach Publisher, Inc.
- Moyer, D. David and Kenneth Paul Fisher. 1973. Land Parcel Identifiers for Information Systems, Chicago, IL: American Bar Association.
- Municipal/County Executive Directory 1989/1990. 1990. Cosal Publication Company, Washington, DC.
- National Research Council. 1980. Need for a Multi-purpose Cadastre, Washington, DC: National Academy Press.
- National Research Council. 1983. Procedures and Standards for a Multipurpose Cadastre, Washington, DC: National Academy Press.
- Nyerges, Timothy. 1989. "TIGER Tales of GIS for Transportation at the University of Washington," Progress Report on TIGER Demonstration Project, Seattle, WA.
- O'Reagan, R. Thomas and Alan J. Saalfeld. 1987. "Geocoding Theory and Practice at the Bureau of the Census," URISA, pp. 231-242.
- Pigot, Simon and Peter Zavart. 1989. "Assigning Attributes Data to Spatial Entities," ESRI Proceedings.
- Pissimissis, Joannis. 1989. "Information System for Urban and Regional Planning: Who is in the Driver's Seat?" URISA, pp. 54-66.

- Prisley, Steven and Roy A. Mead. 1987. "Cost-Benefit Analysis for Geographic Information System," In Proceedings of GIS 1987 Second Annual International Conference, Exhibits and Workshops on Geographic Information Systems, American Society of Photogrammetry and American Congress on Surveying and Mapping, Falls Church, VA, pp. 29-45.
- Reynolds, H. T. 1977. The Analysis of Cross-Classifications, The Free Press, NY.
- Roddy, Alan W. 1988. "GIS in City of Miami: An Emphasis on Database Development," URISA, Vol. II, pp. 57-61.
- Simmons, Al. 1989, August. "Census Bureau Map Tapes Are Ready," Government Technology, Vol. 2, No. 5.
- Simutis, Leonard J. 1980. "Development of Social Indicator and Neighborhood Analysis from Computerized Records," Computers in Local Government, Urban and Regional Planning, Pennsauken, NJ, Auerbach Publisher, Inc.
- Smith, Robert H. and Thomas L. East. 1989. "Automated Mapping and Information Management for Public Utility Services in Mississippi," URISA, Vol. IV, pp. 223-233.
- Smith, Terence R., Sudhakar Menon, Jeffrey L. Star and John Cater. 1987. "Requirements and Principles for Implementation and Construction of Large-scale Geographic Systems," Geographic Information Systems, Vol. 1, No. 1, pp. 13-31.
- Somers, Rebecca. 1987. "Geographic Information Systems in Local Government: A Commentary," Photogrammetric Engineering and Remote Sensing, Vol. 53, No. 10, pp. 1379-1382.
- Somers, R. and P. Eichelberger. 1987. "The Development of an Integrated Cadastral Database," Paper presented to the 1987 Annual Conference of the Urban and Regional Information Systems Association URISA, pp. 75-83.
- SORSA 5. International Colloquium on Street Segment-Oriented Referencing Systems. March, 1978. Bonn, Germany.
- Star, Jeffrey and John Estes. 1990. Geographic Information Systems: An Introduction, Prentice-Hall, Inc.
- Stoner, John E. 1983. "Automated Record Systems for Utilities," Computer Graphics, Vol. 6, No. 7, pp. 51-62.
- System Development Corporation (SDC). May 21, 1969. "A Geographic Base File for Urban Data Systems," Paper presented by SDC's Bob Totsckek, Vald Almemdinger and Ken Needham to Meeting of the Bay Area Automated Information System Coordinating Committee, San Francisco, CA.
- Teicholz, Eric and Brian J. L. Berry. 1983. Computer Graphics and Environmental Planning, Prentice-Hall, Inc., Englewood Cliffs, NJ.



- U. S. Department of Commerce, Bureau of the Census. 1970. Census Use Study No. 4: The DIME Geocoding System, Washington, DC: Superintendent of Documents, U.S. Government Printing Office.
- U. S. Department of Commerce, Bureau of the Census. 1971a. Geographic Base Files: Plans, Progress, and Prospects, Proceedings, Jacksonville, Florida, Superintendent of Documents, U.S. Government Printing Office.
- U. S. Department of Commerce, Bureau of the Census. 1971b. Geographic Base Files: System Uses, Maintenance, Problem Solving, Conference Proceedings, Series GE 60 No. 3, Arlington, Texas, Washington, DC: U.S. Bureau of the Census.
- U. S. Department of Commerce, Bureau of the Census. 1971c. Use of Address Coding Guides in Geographic Coding: Case Studies, Wichita, Kansas, Washington, DC, Superintendent of Documents, U.S. Government Printing Office.
- U. S. Department of Commerce, Bureau of the Census, 1974. Geographic Base (DIME) File System: A Forward Look, Proceedings, Boston, Massachusetts, Washington DC, Superintendent of Documents, U.S. Government Printing Office.
- U. S. Department of Commerce, Bureau of the Census, 1986. State and Metropolitan Area Data Book, U.S. Government Printing Office.
- U. S. Department of Housing and Urban Development (HUD). 1968. Urban and Regional Information Systems: Support for Planning in Metropolitan Areas, Washington, DC: Superintendent of Documents, U.S. Government Printing Office.
- U. S. Department of Housing and Urban Development (HUD). 1971. Municipal Information Systems, The State of the Art in 1970: Urban Information Systems Inter-Agency Committee (USAC), Washington, DC.
- U.S. Geological Survey (USGS). 1985. "Digital Line Graphs From 1:100,000 Scale Maps," Data Uses Guide 2, Reston, VA, U.S. Geological Survey.
- URISA News, "1990 Census TIGER Files Now Available on CD-ROM," Issue 100, July 1989, p. 11.
- White, William O., III. 1989. "Digital Map Conversion: Understand the Basics," American City and County, Vol. 104, pp. 84-90.
- Wilkinson, Leland. 1989. SYSTAT: The System for Statistics, Evanston, IL: SYSTAT, Inc.
- Wood, Stearns J. 1989. "Tacoma Planning Applications and GIS Development," In Proceedings of the User Conference, Palm Springs, CA.

## **APPENDIX A**

### **SURVEY INSTRUMENT AND COVER LETTERS**

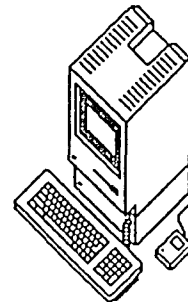
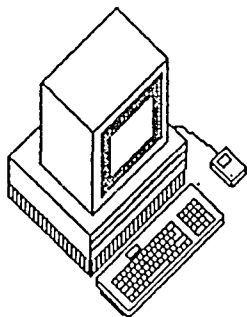
## **GEOCODING SYSTEMS RESEARCH PROJECT**

### **Survey of City Planning Agencies**

**Center for Urban Studies  
Portland State University  
P.O. Box 751  
Portland, OR 97207**

Manual or automated systems have been used by city planning departments/agencies to geocode street addresses to geographic locations. In recent years, a number of those departments/agencies have been turning to advanced systems such as a Geographic Information System (GIS) for this purpose. We are interested in how such advancements affect the generation and use of local data (e.g., employment, housing, travel, or land use). Please answer all questions by circling or filling in the appropriate responses which apply to your situation. Your answers are important to obtain accurate information to assess experience and to draw conclusions about the generation and use of local data in your department/agency.

Thank you for your help with this research.



## METHODS OF GEOCODING, DATA GENERATION, AND USE

The following section asks specific questions concerning methods of geocoding, kinds of data used in planning and their quantity, frequency of update, spatial resolution, and utility.

1. Which of the following geocoding method(s) are used to assign street addresses to geographic locations? (Please circle all that apply.)

### Manual Geocoding

- 1.1 Look up on map
- 1.2 Look up in directory

### Automated Geocoding

- 1.3 Batch processing with tabular output
  - 1.4 Interactive processing with tabular output
  - 1.5 Transfer of geocoded data to a Thematic Mapping Program
  - 1.6 Geocoding within a GIS (e.g. Landtrak, Arc/Info, etc.)
  - 1.7 Geocoding within a specialized system (e.g. car pooling, vehicle dispatching, etc.)
  - 1.8 Other (please specify)
2. Geocoded addresses are applied to many different kinds of applications. These applications can be grouped into categories that suggest a range of data utilization. (Please circle the utilizations in your department.)
- 2.1 Tabulate addresses by statistical area without mapping the result. (e.g., Matching addresses with traffic zones to update employment data or tabulating addresses for mailings to neighborhoods.)
  - 2.2 Aggregating addresses to statistical areas and mapping. (e.g., Assign the location of addresses to area units such as census tracts, zip codes, etc., summarize, and map.)
  - 2.3 Matching addresses to coordinates and/or network nodes. (e.g., Assign the location of addresses to points and map as incidents.)
  - 2.4 Linking coordinate-based addresses to modeling and other application programs. (e.g., Addresses geocoded to point locations and used as inputs to other analytical or modeling processes.)

3. The U.S. Census of Population and Housing is an important source of secondary data for use by planning departments in conjunction with geocoded local data. Which method(s) best describe your use of Census data? (Circle all that apply.)

- 3.1 Extraction of tabular data from printed reports.
- 3.2 Manual mapping of tract data.
- 3.3 Computer mapping of tract data.
- 3.4 Tabulations of block and block group data from summary tapes.
- 3.5 Mapping of block and block group data from summary tapes.
- 3.6 Development of usage rates with geocoded data as the numerator and census data as the denominator.

**IF YOUR DEPARTMENT/AGENCY DOES NOT USE AUTOMATED GEOCODING SYSTEMS,**



**PLEASE GO TO QUESTION 9.  
USERS OF AUTOMATED GEOCODING  
PROCEED**

4. Local data are geocoded to specific locations or areal units. To what areas do you geocode local data? (Please circle all that apply and provide the quantity.)

EXAMPLE	Areal Units	Quantity
①	Planning Areas	25
②	Census Tracts	100
③	Traffic Zones	150

Areal Units (circle all that apply)	Quantity (number of areal units in your jurisdiction)
4.1 Planning Areas	_____
4.2 Census Tracts	_____
4.3 Traffic Zones	_____
4.4 City Block Groups	_____
4.5 City Blocks	_____
4.6 Ownership Parcels	_____
4.7 Other (Please Specify)	_____

5. What type(s) of local data are geocoded in your department? (Please circle all that apply and provide the quantity and frequency of update.)

EXAMPLE		Quantity of Transaction	Frequency of Update
Type of Data			
① Vital Statistics		200/mo.	yearly
② Housing Data		25/day	2/year
⑩ Land Parcel Data		10/day	2 years

Type of Data	Quantity of Transaction	Frequency of Update
5.1 Vital Statistics	_____	_____
5.2 Housing Inspections	_____	_____
5.3 Building Permits/Inspections	_____	_____
5.4 Retail Sales Data	_____	_____
5.7 Employer/Employee Locations	_____	_____
5.8 Auto Registration Data	_____	_____
5.9 Travel Data	_____	_____
5.10 Land Parcel Data	_____	_____
Other Land Data (Please specify type below):		
5.11 _____	_____	_____
5.12 _____	_____	_____

6. What is the digital form of the geographic base file for geocoding? (Please circle all that apply.)

- 6.1 Directory of street names and address ranges and associated areal unit codes.
- 6.2 Cross street index.
- 6.3 Segment based files (e.g., DIME/TIGER).
- 6.4 Address library (universe of all addresses in your jurisdiction).
- 6.5 Others (please specify). \_\_\_\_\_

7. What best describes geographic base file maintenance in your area? (Circle one number.)

- 7.1 Each agency/department maintains their own files.
- 7.2 Shared maintenance within cities/counties.
- 7.3 Shared maintenance metropolitan wide within functions (e.g., public safety, schools, planning, etc.).
- 7.4 Shared maintenance metropolitan-wide for all or most functions.

8. What is the digital map base for display of geocoded data. (Circle all that apply.)

- 8.1 Not Mapped.
- 8.2 Tract/Neighborhood Boundaries.
- 8.3 DIME/TIGER/Street Centerline File.
- 8.4 Scanned map image background with TIGER database.
- 8.5 Vector parcel layer with street rights-of-way.
- 8.6 Other (please describe). \_\_\_\_\_

### COMPUTING AND DEPARTMENT ENVIRONMENTS

The development of computerized systems in a planning department is usually an evolutionary process. The following section asks specific questions concerning the type and configuration of systems, staff, training, and organization in your city.

9. What computer hardware is currently available for use in your department? (Please circle all that currently apply and provide year and number of units.)

	Year Introduced	# Of Seats
9.1 Terminals to a main-frame in a DP unit.	19 _____	_____
9.2 Terminals to a mini-computer in another department	19 _____	_____
9.3 Terminals to a mini-computer in Planning department	19 _____	_____
9.4 Networked Workstations	19 _____	_____
9.5 Microcomputers	19 _____	_____

10. What type(s) of software is currently available in your department. (Please circle all that apply and provide the date introduced.)

	Year Introduced
10.1 Office Automation (Wordprocessing, Spreadsheets, etc.)	19 _____
10.2 Geocoding Systems (Address Matching, Incident Mapping, etc.)	19 _____
10.3 Thematic Mapping (Atlas*Graphics, PC Map, etc.)	19 _____
10.4 Computer-Aided Design and Drafting (AutoCAD, Intergraph IGDS, etc.)	19 _____
10.5 Geographic Information Systems (ARC/INFO, Intergraph MGE, etc.)	19 _____

<p><b>11. How many employees currently work with geocoding and GIS in your city planning agency? (Please circle all that apply and probable FTE.)</b></p> <p style="text-align: center;">Number (full-time equivalent)</p> <p>11.1 Users/Analysts _____</p> <p>11.2 Systems Support _____</p> <p>11.3 Database Administrators _____</p> <p>11.4 Operators/Technicians _____</p> <p>11.5 Others (please specify) _____</p> <hr/> <p><b>12. How have your staff been trained to use geocoding and GIS? (Please circle all that apply and indicate the number of person hours of training in the last year.)</b></p> <p style="text-align: center;">Hours of Training</p> <p>12.1 In-House Training _____</p> <p>12.2 Consultant Training _____</p> <p>12.3 Short Courses _____</p> <p>12.4 Community College/University courses _____</p> <p>12.5 Vendor Training _____</p> <p>12.6 Other (please specify) _____</p> <hr/> <p><b>13. Is the automated system or GIS centralized in your City? (Please circle all that apply.)</b></p> <p>If Yes, which of the following departments utilize GIS services provided by a central unit:</p> <p>13.1 Public Works _____</p> <p>13.2 Data Processing _____</p> <p>13.3 Public Safety _____</p> <p>13.4 Planning _____</p> <p>13.5 Transportation/Traffic _____</p> <p>13.6 Others (please specify) _____</p> <p style="text-align: center;">(In which unit is the centralized GIS located)</p> <p>If No, which of the following departments have a separate GIS function/unit:</p> <p>13.7 Public Works _____</p> <p>13.8 Data Processing _____</p> <p>13.9 Public Safety _____</p> <p>13.10 Planning _____</p> <p>13.11 Transportation/Traffic _____</p> <p>13.12 Others (please specify) _____</p>	<p><b>14. Which of the following organizations in your metropolitan area use an automated system or GIS to geocode street addresses to areal locations? (Please circle all that apply.)</b></p> <p>14.1 None _____</p> <p>14.2 City _____</p> <p>14.3 County _____</p> <p>14.4 Metropolitan Planning Agency _____</p> <p>14.5 Other (Please Specify) _____</p> <p><b>15. Approximately what is the 1990/91 budget of your city planning department/agency?</b></p> <p>15.1 \$ _____</p> <p style="text-align: center; font-weight: bold; margin-top: 20px;">Thank You For Your Assistance.</p> <div style="border: 1px solid black; padding: 10px; text-align: center; margin-top: 20px;"> <p><b>Please Staple Your Business Card Here</b></p> <p><b>We may need to clarify some issues.</b></p> </div> <p style="text-align: center; margin-top: 20px;"><i>or provide the information below</i></p> <table border="1" style="width: 100%; border-collapse: collapse; margin-top: 10px;"> <tr><td style="height: 20px;"> </td></tr> <tr><td style="text-align: center;">Name of Respondent</td></tr> <tr><td style="text-align: center;">Title</td></tr> <tr><td style="text-align: center;">Department of</td></tr> <tr><td style="text-align: center;">Address</td></tr> <tr> <td style="text-align: center;">City</td> <td style="text-align: center;">State</td> <td style="text-align: center;">Zip</td> </tr> <tr> <td style="text-align: center;">Date</td> <td colspan="2" style="text-align: center;">(      ) - Day Time Phone Number</td> </tr> </table>		Name of Respondent	Title	Department of	Address	City	State	Zip	Date	(      ) - Day Time Phone Number	
Name of Respondent												
Title												
Department of												
Address												
City	State	Zip										
Date	(      ) - Day Time Phone Number											

December 14, 1990

NAME  
ADDRESS  
CITY, STATE ZIP

Dear NAME:

City planners often maintain and use detailed geographic data to evaluate planning alternatives. This has required converting or geocoding data referenced by street address to geographic locations. Automated geocoding is becoming prevalent as departments incorporate local records into Geographic Information Systems (GIS).

The Center for Urban Studies at Portland State University is conducting research on the use of geocoding and GIS in city planning departments. This research focuses on the use of geocoding to generate data for use in a GIS. A survey of city planning departments will help us to determine the kinds of data being used and level of detail.

A questionnaire is enclosed. Would you, or a person in your department most familiar with GIS and geocoding, complete and return the questionnaire in the enclosed pre-paid envelope. Approximately 15 to 20 minutes will be required to complete it. The findings of this study will help planners when acquiring systems and will help vendors in developing systems that are responsive to planning needs. If you have any question, please call me on (503) 725 4020.

Thank you for your assistance.

Sincerely,

Fawzi S. Kubbara  
Project Director

FK:wlm

enclosure

December 14, 1990

NAME  
ADDRESS  
CITY, STATE ZIP

Dear NAME:

We seek your help in the pilot test of a survey instrument directed to city planning organizations. Our purpose is to develop an understanding of implementation and uses of computerized systems for geocoding and geographic information.

In addition to completing the survey, we ask that you provide comments about the instrument. We want to make sure it works. Please add your comments on the instrument itself and/or attached sheets.

Thank you for your assistance.

Sincerely,

Kenneth J. Dueker  
Director  
Center for Urban Studies

KJD:fk

enclosure



January 2, 1991

NAME  
ADDRESS  
CITY, STATE ZIP

Dear NAME:

Two weeks ago, a questionnaire seeking information about the effects of Geographic Information Systems (GIS) on local data generation and use was mailed to you. Your agency was one of the planning departments in the pilot study we surveyed.

If you have already completed and returned it to us please accept our sincere thanks. If not, please do so today. Since it has been sent to only a small sample of local planning agencies, it is extremely important that it be tested carefully before conducting the main survey.

If by some chance you did not receive the questionnaire, or it got misplaced, please call me on (503) 725-4020 and a replacement will be mailed to you.

Sincerely

Fawzi S. Kubbara  
Project Director

fsk

February 4, 1991

NAME  
ADDRESS  
CITY, STATE ZIP

Dear NAME:

City planners often maintain and use detailed geographic data to evaluate planning alternatives. This has required converting or geocoding data referenced by street address to geographic locations. Automated geocoding is becoming prevalent as departments incorporate local records into Geographic Information Systems (GIS).

The Center for Urban Studies at Portland State University is conducting research on the use of geocoding and GIS in city planning departments. This research focuses on the use of geocoding to generate data for use in a GIS. A survey of city planning departments will help us to determine the kinds of data being used and level of detail.

A questionnaire is enclosed. Would you, or a person in your department most familiar with GIS and geocoding, complete and return the questionnaire in the enclosed pre-paid envelope. Approximately 15 to 20 minutes will be required to complete it. The findings of this study will help planners when acquiring systems and will help vendors in developing systems that are responsive to planning needs. If you have any question, please call me on (503) 725 4020.

Thank you for your assistance.

Sincerely,

Fawzi S. Kubbara  
Project Director

FK:wlm

enclosure

February 18, 1991

NAME  
ADDRESS  
CITY, STATE ZIP

Dear NAME:

You may remember receiving a questionnaire from the Center for Urban Studies at Portland State University as part of our research to determine the effect of automated geocoding and Geographic Information Systems (GIS) on data generation and use in local planning departments. The questionnaire (cover page enclosed) seeking information for this project was mailed to you about two weeks ago.

As stated in the original cover letter, you or a person familiar with geocoding and GIS are to complete the questionnaire. If you have not done so, I would greatly appreciate if you or the designated person complete and return the questionnaire as soon as possible. If you did not receive the questionnaire, or it has been misplaced, please call me at (503) 725-4020 and a replacement will be mailed to you.

The sample for this project is small, since only planning departments in U.S. cities with 50,000 or more are surveyed, and some of these departments do not utilize automated geocoding.

Many thanks for your help and cooperation in this most important project. I will be glad to provide you with the summary of this project upon your request.

Sincerely

Fawzi S. Kubbara  
Project Director

FK:wlm

March 25, 1991

NAME  
ADDRESS  
CITY, STATE ZIP

Dear NAME:

Local planning departments are under increasing pressure of conducting more comprehensive and detailed plans that require accurate and timely data. As a result, many departments are turning to advanced computer technology such as automated geocoding and Geographic Information Systems (GIS) for this purpose.

The Center for Urban Studies at Portland State University is conducting research on the effect of automated geocoding and GIS on data generation and use in local planning departments. About six weeks ago, you were sent a questionnaire seeking information on the availability and utility of such technology in your department. As of today, we have not received your completed questionnaire.

I am writing to you again because of the significance each questionnaire has to be usefulness of this study. Even if your department has not yet used automated geocoding or GIS, your answers to the other questions are important for this study. Your name was drawn from the Municipal/County Executive Directory 1989/1990. It is important that we have a high response rate.

As I mentioned in our last letter, the questionnaire should be completed by a person who is familiar with the development, use, and costs of the system. In the event that your questionnaire has been misplaced, a replacement is enclosed.

Many thanks for your help and cooperation in this most important project. I will be glad to provide you with the summary of this project upon your request. Results should be available by mid-summer. If you have any question, please call me at (503) 725-4020.

Sincerely,

Fawzi S. Kubbara  
Project Director

FK:wlm

## APPENDIX B

### SUMMARY OF STATISTICS

TABLE B-I  
CITY POPULATION AND AREA MEANS

	Mean Population	Mean Area
All Cities	182,852	61.8 sq. mi.
Respondent Cities	174,724	60.0 sq. mi.
Automated Cities	248,922	80.0 sq. mi.

TABLE B-II  
STATISTICS OF CLUSTERS

VARIABLES	Non-Computer-Oriented Depts.	CADD Mapping Depts.	Using Geocoding on Mainframe	Large Budget in Large-Area Cities	New to GIS and Geocoding Depts.
AREA					
No. of Cases	144	5	58	10	39
Minimum	5.8	16.5	7.1	239.7	3.9
Maximum	217.5	169.2	258.5	759.7	199.4
Mean	39.4	69.6	53.6	435.4	62.6
DECENTRALIZED					
No. of Cases	152	5	58	10	41
Minimum	0.0	0.0	0.0	0.0	1.0
Maximum	6.0	1.0	2.0	6.0	6.0
Mean	0.2	0.4	0.2	1.4	2.6
INCOMEX					
No. of Cases	152	5	58	10	41
Minimum	0.0	0.0	0.0	1.0	0.0
Maximum	21.0	16.0	36.0	16.0	16.0
Mean	3.6	10.6	10.8	7.7	3.1
GEOMETHODS					
No. of Cases	144	5	56	10	38
Minimum	1.0	1.0	1.0	1.0	1.0
Maximum	3.0	3.0	3.0	3.0	3.0
Mean	1.2	2.2	2.3	2.1	2.3
POPULATION					
No. of Cases	152	5	58	10	41
Minimum	47240	55097	49826	52141	50008
Maximum	369365	782248	628088	3485398	736014
Mean	105124	328068	175975	866923	217947
OFFICE					
No. of Cases	127	5	51	10	40
Minimum	1.0	8.0	4.0	4.0	3.0
Maximum	12.0	16.0	14.0	11.0	13.0
Mean	5.7	12.6	9.1	6.5	7.7
CENTRALIZED					
No. of Cases	152	5	58	10	41
Minimum	0.0	0.0	0.0	0.0	0.0
Maximum	6.0	5.0	6.0	6.0	6.0
Mean	0.6	2.0	2.8	1.9	0.4

TABLE B-II  
STATISTICS OF CLUSTERS  
(Continued)

EXCOMEX					
No. of Cases	152	5	58	10	41
Minimum	0.0	10.0	0.0	0.0	0.0
Maximum	7.0	16.0	17.0	10.0	11.0
Mean	0.9	13.2	1.3	1.9	2.8
MICRO					
No. of Cases	97	4	45	10	27
Minimum	1.0	3.0	3.0	1.0	1.0
Maximum	12.0	8.0	17.0	8.0	13.0
Mean	4.8	5.7	8.2	5.3	6.1
DATAUSE					
No. of Cases	95	5	49	9	35
Minimum	1.0	1.0	1.0	1.0	1.0
Maximum	4.0	4.0	4.0	4.0	4.0
Mean	1.9	2.2	2.8	2.6	3.0
CADD					
No. of Cases	35	2	29	4	20
Minimum	1.0	12.0	2.0	1.0	2.0
Maximum	8.0	19.0	23.0	5.0	8.0
Mean	2.9	15.5	6.4	3.2	3.9
GIS					
No. of Cases	22	3	35	7	24
Minimum	1.0	3.0	1.0	1.0	1.0
Maximum	22.0	13.0	13.0	7.0	5.0
Mean	4.1	9.3	4.1	3.0	2.5
THEMATIC					
No. of Cases	13	3	17	6	18
Minimum	1.0	5.0	2.0	1.0	2.0
Maximum	18.0	13.0	17.0	5.0	11.0
Mean	5.1	9.0	7.5	3.0	4.6
GEOCODING					
No. of Cases	17	4	22	4	19
Minimum	1.0	2.0	2.0	1.0	2.0
Maximum	18.0	7.0	21.0	23.0	14.0
Mean	4.8	3.5	10.0	8.7	4.5
BUDGET					
No. of Cases	124	2	50	9	33
Minimum	70524	223000	170000	341208	130000
Maximum	10000000	895570	5400000	48000000	6100000
Mean	941349	559285	1129685	8272621	1498959



## **APPENDIX C**

### **CITIES SURVEYED**

\* Responding cities with Geographic Information Systems in planning departments.

Anchorage	AK	Irvine*	CA	Torrance	CA
Birmingham*	AL	La Mesa	CA	Vallejo	CA
Huntsville*	AL	Lakewood	CA	Ventura*	CA
Mobile	AL	Long Beach*	CA	Walnut Creek	CA
Montgomery	AL	Los Angeles*	CA	West Covina	CA
Tuscaloosa*	AL	Modesto	CA	Westminster	CA
Fort Smith	AR	Montebello	CA	Whittier	CA
Little Rock	AR	Monterey Park	CA	Arvada	CO
North Little Rock	AR	Mountain View	CA	Aurora	CO
Pine Bluff	AR	Napa	CA	Boulder	CO
Glendale	AZ	Newport Beach	CA	Colorado Springs	CO
Mesa*	AZ	Norwalk	CA	Denver	CO
Phoenix*	AZ	Oakland*	CA	Fort Collins	CO
Scottsdale	AZ	Oceanside	CA	Greeley	CO
Tempe*	AZ	Ontario*	CA	Lakewood	CO
Tucson*	AZ	Orange	CA	Longmont	CO
Alameda	CA	Oxnard	CA	Pueblo*	CO
Alhambra	CA	Palo Alto	CA	Westminster	CO
Anaheim	CA	Pasadena	CA	Bridgeport	CT
Bakersfield	CA	Pico Rivera	CA	Bristol	CT
Baldwin Park	CA	Pomona	CA	Danbury	CT
Bellflower	CA	Rancho Cucamonga	CA	East Hartford	CT
Berkeley	CA	Redding	CA	Fairfield	CT
Buena Park	CA	Redondo Beach	CA	Greenwich	CT
Burbank	CA	Redwood City*	CA	Hamden	CT
Carson	CA	Richmond	CA	Hartford	CT
Cerritos	CA	Riverside	CA	Meriden	CT
Chula Vista	CA	Sacramento	CA	New Britain	CT
Compton	CA	Salinas	CA	New Haven	CT
Concord	CA	San Bernardino*	CA	Norwalk	CT
Costa Mesa	CA	San Diego*	CA	Stamford*	CT
Daly City	CA	San Francisco	CA	Stratford	CT
Downey	CA	San Jose	CA	Waterbury	CT
El Monte	CA	San Leandro	CA	West Haven	CT
Escondido	CA	San Mateo	CA	Washington	DC
Fountain Valley	CA	Santa Ana	CA	Wilmington	DE
Fremont	CA	Santa Barbara	CA	Boca Raton	FL
Fresno	CA	Santa Clara	CA	Cape Coral	FL
Fullerton	CA	Santa Monica	CA	Clearwater	FL
Garden Grove	CA	Santa Rosa	CA	Daytona Beach	FL
Glendale*	CA	Simi Valley	CA	Fort Lauderdale	FL
Hawthorne	CA	South Gate	CA	Gainesville*	FL
Hayward	CA	Stockton	CA	Hialeah	FL
Huntington Beach	CA	Sunnyvale	CA	Hollywood	FL
		Thousand Oaks	CA	Jacksonville	FL

Lakeland	FL	South Bend	IN	Southfield	MI
Largo	FL	Terre Haute	IN	St Clair Shores	MI
Melbourne	FL	Kansas City	KS	Sterling Heights	MI
Miami*	FL	Lawrence	KS	Taylor	MI
Miami Beach*	FL	Olathe	KS	Troy	MI
Orlando	FL	Overland Park*	KS	Warren	MI
Pensacola	FL	Topeka	KS	Westland	MI
Pompano Beach	FL	Wichita	KS	Wyoming	MI
Sarasota	FL	Louisville	KY	Bloomington	MN
St Petersburg	FL	Owensboro	KY	Duluth	MN
Tallahassee	FL	Baton Rouge	LA	Minneapolis*	MN
Tampa	FL	Kenner	LA	St Paul	MN
West Palm Beach	FL	Lafayette	LA	Columbia	MO
Albany*	GA	Lake Charles	LA	Florissant	MO
Atlanta	GA	Monroe	LA	Independence	MO
Columbus	GA	New Orleans*	LA	Kansas City	MO
Macon	GA	Shreveport	LA	Springfield*	MO
Savannah*	GA	Boston*	MA	St Joseph*	MO
Honolulu*	HI	Brockton	MA	St Louis	MO
Cedar Rapids	IA	Cambridge	MA	Jackson*	MS
Council Bluffs	IA	Chicopee	MA	Billings	MT
Davenport	IA	Fall River*	MA	Great Falls	MT
Des Moines	IA	Framingham	MA	Asheville	NC
Dubuque	IA	Lawrence	MA	Charlotte	NC
Iowa City	IA	Lowell	MA	Durham	NC
Sioux City	IA	New Bedford	MA	Gastonia	NC
Waterloo	IA	Pittsfield	MA	Greensboro	NC
Boise	ID	Somerville*	MA	High Point	NC
Arlington Heights	IL	Springfield	MA	Raleigh	NC
Aurora	IL	Weymouth	MA	Winston-Salem	NC
Champaign	IL	Worcester	MA	Fargo	ND
Chicago	IL	Baltimore*	MD	Lincoln	NE
Decatur	IL	Portland	ME	Omaha	NE
Des Plaines	IL	Ann Arbor*	MI	Manchester	NH
Elgin*	IL	Battle Creek	MI	Nashua	NH
Evanston	IL	Dearborn	MI	Bayonne	NJ
Joliet	IL	Detroit*	MI	Camden	NJ
Mount Prospect	IL	East Lansing	MI	Cherry Hill	NJ
Oak Park	IL	Flint	MI	Elizabeth	NJ
Peoria	IL	Grand Rapids	MI	Hamilton	NJ
Rockford	IL	Kalamazoo	MI	Jersey City	NJ
Schaumburg	IL	Lansing	MI	Middletown	NJ
Skokie	IL	Livonia	MI	Toms River	NJ
Waukegan	IL	Pontiac	MI	Trenton	NJ
Fort Wayne	IN	Roseville	MI	Union	NJ
Indianapolis*	IN	Royal Oak	MI	Vineland	NJ
Muncie	IN	Saginaw	MI	Woodbridge	NJ

Albuquerque*	NM	Oklahoma City	OK	Galveston	TX
Las Cruces*	NM	Tulsa*	OK	Garland	TX
Santa Fe	NM	Eugene*	OR	Houston	TX
Las Vegas	NV	Portland*	OR	Irving*	TX
Reno	NV	Salem*	OR	Killeen	TX
Albany	NY	Abington	PA	Laredo	TX
Binghamton	NY	Allentown	PA	Longview*	TX
Buffalo	NY	Ardmore	PA	Lubbock	TX
Clay	NY	Bensalem	PA	McAllen	TX
Elmsford	NY	Bethlehem	PA	Mesquite	TX
Endwell	NY	Harrisburg	PA	Midland	TX
Huntington	NY	Lancaster	PA	Odessa	TX
Islip	NY	Penn Hills	PA	Pasadena	TX
Lindenhurst	NY	Philadelphia	PA	Plano	TX
Manhasset	NY	Pittsburgh*	PA	Port Arthur	TX
Mount Vernon	NY	Reading	PA	Richardson	TX
New City	NY	Scranton	PA	San Angelo	TX
New Rochelle	NY	Wilkins-Barre	PA	San Antonio	TX
New York	NY	Cranston	RI	Tyler	TX
Newtonville	NY	East Providence	RI	Waco	TX
Niagra Falls*	NY	Pawtucket	RI	Wichita Falls	TX
Oyster Bay	NY	Providence	RI	Ogden	UT
Rochester*	NY	Warwick	RI	Orem*	UT
Rochester	NY	Charleston	SC	Provo	UT
Schenectady	NY	Columbia	SC	Salt Lake City	UT
Smithtown	NY	Greenville	SC	Sandy City*	UT
Syracuse	NY	North Charleston	SC	West Valley City	UT
Troy	NY	Rapid City	SD	Alexandria	VA
West Seneca	NY	Sioux Falls	SD	Chesapeake	VA
Williamsville	NY	Chattanooga	TN	Hampton	VA
Akron	OH	Clarksville	TN	Lynchburg	VA
Canton	OH	Knoxville	TN	Newport News*	VA
Cincinnati	OH	Memphis	TN	Norfolk*	VA
Cleveland	OH	Nashville*	TN	Portsmouth	VA
Cleveland Heights	OH	Abilene	TX	Richmond	VA
Columbus	OH	Amarillo	TX	Roanoke	VA
Dayton*	OH	Arlington*	TX	Suffolk	VA
Euclid	OH	Austin	TX	Virginia Beach	VA
Hamilton	OH	Baytown	TX	Bellevue	WA
Kettering	OH	Beaumont	TX	Everett	WA
Lakewood	OH	Brownsville	TX	Seattle	WA
Springfield	OH	Bryan	TX	Spokane	WA
Toledo	OH	Corpus Christi	TX	Tacoma*	WA
Warren	OH	Dallas	TX	Yakima*	WA
Youngstown	OH	Denton	TX	Appleton	WI
Lawton	OK	El Paso	TX	Eau Claire	WI
Norman	OK	Fort Worth	TX	Green Bay	WI

Janesville	WI
Kenosha	WI
Madison*	WI
Milwaukee*	WI
Racine	WI
Waukesha	WI
Wauwatosa	WI
West Allis	WI
Charleston	WV
Huntington	WV
Cheyenne	WY